APPLICATION OF NEW TECHNOLOGY FOR THE DIAGNOSIS OF VIRAL INFECTION

By

Jayme Parker, B.S., M.S.P.H.

A Dissertation Submitted in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

In

Biological Sciences

University of Alaska Fairbanks

August 2019

APPROVED:

Dr. Jack Chen, Committee Chair
Dr. Karsten Hueffer, Committee Member
Dr. Andrea Ferrante, Committee Member
Dr. Bernd Jilly, Committee Member
Dr. Leah Berman, Dean
College of Natural Science and Mathematics
Dr. Michael Castellini,
Dean of the Graduate School

Abstract

New technology is challenging conventional methods for characterizing pathogenic viruses in clinical laboratories. These newer methods are superior to older methods due to their ability to broadly target numerous pathogens in multiplexed ways. Even more intriguing, new technologies are capable of detecting viruses in non-targeted manners. Before these newer methods can be adopted by accredited medical laboratories, they must be validated to assess whether or not they meet minimum federal standards in terms of assay accuracy, precision, reproducibility, and cross-reactivity. This thesis begins to answer important questions facing clinical laboratories when adopting new technology. In Chapter 1, assays targeting single virus types are compared to a multiplexed assay using a proprietary electrochemical detection technology to determine if multiplexing has a detrimental effect on analytical sensitivity when detecting respiratory viruses simultaneously. Chapter 2 focuses on the issue of false positivity when testing for viruses in low-prevalence populations. To evaluate this, a multiplex flow immunoassay technology is used to perform surveillance of human immunodeficiency virus (HIV) infection in Alaskans, a low HIV-prevalence population. Chapter 3 describes clinical diagnostic applications of next-generation sequencing (NGS) providing examples of how NGS compares to conventional methods for characterizing pathogenic viruses such as hepatitis C virus, herpesvirus, adenovirus, and influenza virus. The final chapter describes how NGS can be used to characterize viruses by geographical region of transmission by analyzing an outbreak of canine parvovirus that occurred in the interior of Alaska. This chapter serves as a clear example of NGS's appeal to enhancing our epidemiological understanding during outbreaks. Although there are significant challenges to implementation, especially for NGS, each chapter shows promise in new technologies for clinical laboratories.

Table of Contents

Page

Title Page i

Abstract iii

Table of Contents v

List of Figures ix

List of Tables xiii

List of Appendices xv

Acknowledgements xvii

General Introduction 1

References 6

Chapter 1 Analytical sensitivity comparison between singleplex real-time PCR and a multiplex PCR platform for detecting respiratory viruses 9

1.1 Abstract 9

1.2 Introduction 10

1.3 Methods and Materials 12

1.3.1 Clinical specimens 12

1.3.2 Preparation of standard materials 12

1.3.3 Determination of singleplex real-time LOD 13

1.3.4 Conversion of $TCID_{50}$/mL concentration to copy number 15

1.4 Results 16

1.5 Conclusion 20

1.6 Acknowledgements 22

1.7 References 26

1.8 Appendices 29

Chapter 2 BioRad BioPlex® HIV Ag-Ab assay: Incidence of false positivity in a low-prevalence population and its effects on the current HIV testing algorithm 67

2.1 Abstract 67

2.2 Introduction 68

2.3 Methods and Materials 69

2.4 Results 69

2.5 Discussion 71

2.6 Ethical Approval 72

2.7 Acknowledgements 72

2.8 References 73

Chapter 3 Application of next generation sequencing for the detection of human viral pathogens in clinical specimens 79

3.1 Abstract 79

3.2 Introduction 80

3.3 Methods and Materials 80

3.3.1 Specimens 80

3.3.2 Construction of sequencing library 81

3.3.3 Sequencing and analysis 81

3.3.4 Conventional Viral Detection Methods 81

3.3.4.1 Viral Culture 82

3.3.4.2 Fluorescent Microscopy 82

3.3.4.3 Serum Neutralization (SN) Assay 83

3.3.4.4 Enzyme-Linked Immunosorbant Assay (ELISA) 83

3.3.4.5 Real-time quantitative PCR (qPCR) 83

3.3.5 Comparative Sequencing Method - Pyrosequencing 84

3.4 Results 85

3.4.1 NGS for detecting clinical adenovirus infections 85

3.4.2 NGS for detecting clinical herpesvirus infections 85

3.4.3 Further characterization of viral hepatitis C and G 87

3.4.4 Antiviral resistance of influenza viruses in clinical specimens 88

3.4.5 Non-specific viral sequencing reads in NGS data 89

3.5 Discussion 90

3.6 Competing Interests 93

3.7 Ethical Approval 93

3.8 Acknowledgements 93

3.9 References 100

Chapter 4 Investigation of canine parvovirus outbreak in Alaska using next generation sequencing.....105

4.1 Abstract 105

4.2 Introduction.....106

4.3 Materials and Methods.....108

4.3.1 Specimens 108

4.3.2 Referral Testing (serology, PCR, & genotyping) 108

4.3.3 Nucleic acid isolation in preparation for sequencing 109

4.3.4 Library preparation for sequencing.....109

4.3.5 Sequencing and analysis 110
4.4 Results 110
4.5 Discussion 112
4.6 Funding 112
4.7 Competing interests 112
4.8 Acknowledgements 113
4.9 References 118
General Conclusion 121
References 125
Appendices 127

List of Figures Page

Figure 1.A-1: Adenovirus (Ad2-4) plasmid details 30

Figure 1.A-2: Influenza A (FluA-7) plasmid details 31

Figure 1.A-3: Influenza A/H3 (H3-4-49) plasmid details 32

Figure 1.A-4: Influenza A/H1N1 (pdmH1-40) plasmid details 33

Figure 1.A-5: Influenza B (FluB-6) plasmid details 34

Figure 1.A-6: Parainfluenza 1 (Para1-22) plasmid details 35

Figure 1.A-7: Parainfluenza 2 (Para2-12) plasmid details 36

Figure 1.A-8: Parainfluenza 3 (Para3-3) plasmid details 37

Figure 1.A-9: Respiratory syncytial virus A (RSVA-1) plasmid details 38

Figure 1.A-10: Respiratory syncytial virus B (RSVB-3) plasmid details 39

Figure 1.A-11: Rhinovirus (Rhino-2) plasmid details 40

Figure 1.B-1: Adenovirus real-time PCR assay limit of detection determination 42

Figure 1.B-2: Influenza A (generic) real-time PCR assay limit of detection determination 43

Figure 1.B-3: Influenza A/H3 real-time PCR assay limit of detection determination 44

Figure 1.B-4: Influenza A/H1N1 real-time PCR assay limit of detection determination 45

Figure 1.B-5: Influenza B real-time PCR assay limit of detection determination 46

Figure 1.B-6: Parainfluenza 1 real-time PCR assay limit of detection determination 47

Figure 1.B-7: Parainfluenza 2 real-time PCR assay limit of detection determination 48

Figure 1.B-8: Parainfluenza 3 real-time PCR assay limit of detection determination 49

Figure 1.B-9: RSVA real-time PCR assay limit of detection determination 50

Figure 1.B-10: RSVB real-time PCR assay limit of detection determination 51

Figure 1.B-11: Rhinovirus real-time PCR assay limit of detection determination 52

Figure 1.C-1: Adenovirus (group C, VR-1) TCID concentration conversion to copy numbers ...54

Figure 1.C-2: Adenovirus (group E, VR-1572) TCID concentration conversion to copy numbers..55

Figure 1.C-3: Influenza A (VR-547) CEID concentration conversion to copy numbers 56

Figure 1.C-4: Influenza A/H3 (VR-547) CEID concentration conversion to copy numbers 57

Figure 1.C-5: Influenza A/H1N1 (VR-1736) TCID concentration conversion to copy numbers .58

Figure 1.C-6: Influenza B (VR-101) CEID concentration conversion to copy numbers 59

Figure 1.C-7: Parainfluenza 1 (VR-94) TCID concentration conversion to copy numbers 60

Figure 1.C-8: Parainfluenza 2 (VR-92) TCID concentration conversion to copy numbers 61

Figure 1.C-9: Parainfluenza 3 (VR-93) TCID concentration conversion to copy numbers 62

Figure 1.C-10: RSVA (VR-1540) TCID concentration conversion to copy numbers 63

Figure 1.C-11: RSVB (VR-955) TCID concentration conversion to copy numbers 64

Figure 1.C-12: Rhinovirus (VR-1540) TCID concentration conversion to copy numbers 65

Figure 2.1: Current recommended HIV testing algorithm showing suggested modifications in red. 77

Figure 3.1: Comparison of NGS and conventional virology assays for detecting adenovirus infection. 96

Figure 3.2: Comparison of NGS and conventional virology assays for detecting herpesvirus infection 97

Figure 3.3: Detection and characterization of hepatitis C and G viruses in 5 different sera using NGS.....98

Figure 3.4: Antiviral resistance characterization of influenza viruses using NGS 99

Figure 4.1: Phylogenetic tree of clinical specimen isolates in relation to reference sequences of canine parvovirus 2a, 2b, and vaccine candidate genomes.....116

Figure 4.2: Visual depiction of sequence alignment to reference genome, VP2 gene location within the canine parvovirus genome and analysis of 18 amino acid positions 117

Figure B.1: Comparison of NGS and conventional virology assays for detecting adenovirus infection [Original Publication]. 147

Figure B.2: Comparison of NGS and conventional virology assays for detecting herpesvirus infection [Original Publication]. 149

Figure B.3: Detection and characterization of hepatitis C and G viruses in 5 different sera using NGS [Original Publication]. 148

Figure B.4: Antiviral resistance characterization of influenza viruses using NGS [Original Publication]. 153

List of Tables Page

Table 0.1: Known human viral pathogens classified by specimen origin used for detection 4

Table 0.2: Minimum parameters to establish acceptability of a new assay for human specimen testing.....5

Table 1.1: Plasmid concentrations and copy number determination 23

Table 1.2: Limit of detection (LOD) comparison summary 24

Table 1.3: Relationship between $TCID_{50}$/mL concentrations and copy number.....25

Table 2.1: Comparison of BioPlex® HIV ag-ab assay IDX values by result type.....75

Table 2.2: Description of false positive results from BioPlex® HIV ag-ab assay 76

Table 3.1: NGS proficiency compared to pyrosequencing methodology for detecting antiviral resistance in influenza A virus 95

Table 4.1: Specimens and referral results summary 114

Table 4.2: NGS metrics after aligning read files to reference genome NC_001539.1 (canine parvovirus) 115

Table A.1: Plasmid concentrations and copy number determination [Original Publication].....131

Table A.2: Limit of detection (LOD) comparison summary [Original Publication].....133

Table A.3: Relationship between $TCID_{50}$/mL concentrations and copy number [Original Publication].....135

Table B.1: NGS proficiency compared to pyrosequencing methodology for detecting antiviral resistance in influenza A virus [Original Publication].....155

Table C.1: Comparison of methods for detecting DNA viruses [Original Publication].....164

Table C.2: NGS for characterizing hepatitis C viruses and potential resistance to infection inhibitors [Original Publication].....165

Table C.3: NGS for characterizing influenza viruses [Original Publication].....166

List of Appendices Page

Appendix 1.A Plasmid construction and quantification details 29

Appendix 1.B Establishing limits of detection for singleplex real-time PCR assays.....41

Appendix 1.C Converting TCID dilutions to copy number equivalents 53

Appendix A. Correction: Analytical Sensitivity Comparison between Singleplex Real-Time PCR and a Multiplex PCR Platform for Detecting Respiratory Viruses [Original Publication] 127

Appendix B. Application of next generation sequencing for the detection of human viral pathogens in clinical specimens [Original Publication].....143

Appendix C. Next generation sequencing in clinical virology diagnostics [Original Publication] 161

Acknowledgements

I would like to thank the staff of the Alaska State Virology Laboratory for allowing me the time and scientific freedom to think deeply about how we are conducting viral surveillance for the State of Alaska and make significant contributions to their future efforts. It is their "behind the scene" actions in the medical world and diligence on the work bench that helps keep Alaskan's safe from preventable infectious disease. I would like to thank my mom, the elementary school teacher, who encourages me to continually educate myself whether I'm enrolled in school, or not. I especially thank my husband who has heroically raised our two young girls throughout this process, whom have endured most of their short lives with their mom staring at this thesis, asking questions, and learning about virology whether they wanted to or not. I also thank Dr. Jack Chen, who took a chance on me and offered me the research assistantship necessary for me to even consider such a project. I look forward to using this degree to make a difference in public health laboratory diagnostics in the U.S. and abroad.

General Introduction

Clinical laboratories continually monitor the effectiveness of their assays to detect novel and variant pathogens and openly recognize the limitations of target-dependent testing. For this reason, new technologies for the diagnosis and monitoring of infectious pathogens are increasingly being utilized to account for the breadth of microbial diversity in clinical specimens [1]. Whole genome sequencing technology and its ability for broad characterization of all genetic information within a clinical specimen has helped enhance, as well as complicate, infectious disease analyses in the laboratory.

Commercial next-generation sequencing (NGS) platforms were launched in 2005 [2], a technology that has challenged the Sanger sequencing platforms which served as the gold standard for sequencing for over 30 years. In 2009, this newer sequencing technology was described in the literature as a potential tool not only for research, but for clinical diagnostics as well [3]. Now, nine years later, we are only beginning the validation processes necessary to bring sequence-based assays into the infectious disease clinical laboratory. The barriers preventing faster implementation include high start-up costs, confusion surrounding the clinical interpretation of results, lack of analysis pipelines (i.e. bioinformatics), and data storage [4]. This thesis contributes to the validation of assays utilizing new technology, including NGS, for the purposes of viral pathogen detection to address these issues.

Interpretation of sequencing data in clinical specimens is complicated by the large quantities of diverse genetic material contained in various matrices. For instance, trace amounts of pathogen nucleic acid can be found in clinical specimens but may not necessarily be attributable to the cause of disease. Although much research has been conducted describing the metagenomics of human biological materials, this has largely been centered on the prokaryotic

fraction of our microbiome [5]. In this thesis, clinical specimens derived from blood, skin swabs, the gastrointestinal tract, and the respiratory tract are evaluated for pathogenic viruses amongst the background of their metagenomes.

Viruses as a group lack a universal genetic biomarker, like bacteria's 16S ribosomal RNA, and have high mutation rates across small genomes making it difficult to design robust, long-lasting targeted assays [6]. The presence of known viral pathogens in different specimen matrices are briefly described in Table 0.1 to showcase the challenge of designing individual targeted tests for various human infections based on the sheer number of possible targets. Metagenomics is proven to be a powerful tool with endless areas of applications in virology, including the ability to identify novel or variant viruses in a target-independent manner [7]. However, using NGS for clinical diagnostics in virology can be complicated due to the presence of contaminating viruses in the testing environment and the human endogenous retroviruses within the human genome itself [8].

Federal regulations put in place to protect patients from receiving uninterpretable or inaccurate laboratory results require clinical laboratories to perform comprehensive validations to adopt assays that are not approved by the Federal Drug Administration (FDA). To comply with regulations, the performance of an assay must be documented and deemed acceptable by the clinical laboratory performing the test prior to releasing results to a patient. Acceptability is dependent upon specific parameters that must be established for each assay (Table 0.2). Since target-independent assays are capable of detecting novel viruses as well as potentially all variations, it is not feasible to validate all possible outcomes and therefore a combination of a "methods-based" and "analyte-specific" validation approach can be used [9, 10]. Despite

challenges, validation protocols have been described for NGS assays [9, 11-14], and clinical applications are continuing to be described [15-17].

In this thesis, I compare conventional virus detection methods to those comprised of new technologies capable of broad pathogen detection. I also describe how data obtained by newer technologies can be applied to improving our overall epidemiological understanding of virus transmission during an outbreak scenario. Chapter 1 demonstrates a validation method in which to compare the analytical sensitivity of a new commercial electrochemical detection platform to commonly used real-time PCR assays for detecting large panels of respiratory viruses. Chapter 2 evaluates the positive predictive value of a new HIV testing platform using multiplex flow immunoassay technology when performing surveillance in a population with low-HIV prevalence. Chapter 3 contrasts conventional virus detection strategies with NGS in terms of turnaround time and result comparability. The sequencing data obtained from the experiment described in chapter 3 reveals the true scarcity of targeted viral nucleic acid among the total nucleic acid captured from a clinical specimen. However, despite minimal representation in the specimen, clinically relevant information from targeted viral genomes pertaining to disease prophylaxis or treatment, such as the presence of antiviral resistance markers, is shown to be achieved in the initial non-targeted sequencing of a clinical specimen. Chapter 4 describes an outbreak investigation of canine parvovirus based on NGS data recovered from rectal swabs which describes how NGS data can be used to enhance our epidemiological understanding of disease transmission during an outbreak.

Table 0.1: Known human viral pathogens classified by specimen origin used for detection

Specimen origin	Viral pathogens of known clinical importance
GI tract (stool, rectal swabs)	norovirus (3 genogroups), rotavirus (40 strains), enterovirus (>60 serotypes, including polio) , hepatitis A & E virus, adenovirus (7 species, >50 types), sapovirus (4 genogroups), astrovirus (8 species), picobirnavirus, parechovirus (6 types), influenza (3 types), coronavirus (6 types), torovirus [18, 19]
Blood (serum, plasma, buffy coat)	Herpesviruses (8 types), anelloviruses (3 genera), parvovirus B19, human papillomavirus, adenovirus (7 species, >50 serotypes), HIV (2 types), human polyomavirus (11 types), hepatitis B virus, hepatitis C virus (6 genotypes), influenza A (2 types) [20]
Respiratory tract (nasal swabs, pharyngeal swabs, sputum, aspirates, washes)	picornaviruses (rhinovirus (>100 serotypes), enterovirus (>60 serotypes), parechoviruses), paramyxoviruses (respiratory syncytial virus (2 groups), parainfluenzavirus (4 types), metapneumovirus, measles), influenza (3 types), coronaviruses (4 types), adenovirus (5 serotypes), parvovirus, herpesvirus (4 species), anelloviruses (3 genera), papillomaviruses (>100 types), polyomaviruses (11 types) [21]
Skin (rashes, lesions)	Herpesviruses (8 types), Polyomavirus (11 types), Molluscum contagiosum, human papillomavirus (>100 types) [22-25]

Table 0.2: Minimum parameters to establish acceptability of a new assay for human specimen testing

Parameter	Protocol for targeted assays	Adaptation for NGS (non-targeted)
Accuracy	Test x number of specimens simultaneously using the gold standard method and the new assay. [26] Count the number of: 1) True Positives (TP): target is present on both assays 2) True Negatives (TN): target is absent on both assays 3) False Positive (FP): target is present on the new assay only 4) False Negative (FN): target is absent on the new assay only Diagnostic Sensitivity = TP/(TP+FN) Diagnostic Specificity = TN/(TN+FP) Positive Predictive Value = TP/(TP+FP) Negative Predictive Value = TN/(TN+FN)	Compare outcomes of NGS to those obtained by preferred (gold standard) methods, even if methodologies are different [27]. And Compare the obtained sequences to the accepted reference sequence and evaluate depth of coverage (X), allelic read percentage, quality scores, and degree of coverage [11].
Precision	Test the same positive and negative specimens repeatedly on the same run, on different runs, and by different people. Intra-assay precision: Calculate coefficient of variability of values obtained from duplicates on the same run Inter-assay precision: Calculate coefficient of variability of values obtained from duplicates on different runs [28]	Prepare a specimen in duplicate for testing on the target-independent system. Have a different person repeat the duplicates on a different day. [11, 12]
Reportable Range	Report the span of test result values for which the test system is accurate [29]	Define the portion of the genome for which sequence information can be reliably derived. [11]
Reference Range	Define the range of test values expected for the normal healthy population [29]	Normal variation of the sequence of interest within the population that the assay is designed to test. [11]
Analytical Sensitivity	Perform replicates of multiple dilutions of a specimen with known analyte quantity to determine the lowest reliably detectable concentration. Also called the limit of detection (LOD). [30]	Perform replicates of specimens containing known disease-causing variants, calculate the average FN rate [11, 12]
Analytical Specificity	Spike specimens with genetically similar organisms and specimen-related interfering substances (hemolysis, lipemia, medications, etc.) to determine cross-reactivity [30]	Perform replicates of specimens <u>not</u> containing known disease-causing variants, calculate the average FP rate [12]

References

1. Lefterova, M.I., et al., *Next-Generation Sequencing for Infectious Disease Diagnosis and Management: A Report of the Association for Molecular Pathology.* The Journal of Molecular Diagnostics, 2015. **17**(6): p. 623-634.

2. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. 437: p. 376.

3. Voelkerding, K.V., S.A. Dames, and J.D. Durtschi, *Next-Generation Sequencing: From Basic Research to Diagnostics.* Clinical Chemistry, 2009. **55**(4): p. 641-658.

4. Biesecker, L.G., *Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: lessons from the ClinSeq project.* Genet Med, 2012. **14**(4): p. 393-398.

5. Lloyd-Price, J., G. Abu-Ali, and C. Huttenhower, *The healthy human microbiome.* Genome Medicine, 2016. **8**(1): p. 51.

6. Reyes, A., et al., *Going viral: next generation sequencing applied to human gut phage populations.* Nature reviews. Microbiology, 2012. **10**(9): p. 607-617.

7. Mokili, J.L., F. Rohwer, and B.E. Dutilh, *Metagenomics and future perspectives in virus discovery.* Curr Opin Virol, 2012. **2**(1): p. 63-77.

8. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

9. Aziz, N., et al., *College of American Pathologists' laboratory standards for next-generation sequencing clinical tests.* Arch Pathol Lab Med, 2015. **139**(4): p. 481-93.

10. Schrijver, I., et al., *Opportunities and challenges associated with clinical diagnostic genome sequencing: a report of the Association for Molecular Pathology.* J Mol Diagn, 2012. **14**(6): p. 525-40.

11. Gargis, A.S., L. Kalman, and I.M. Lubin, *Assuring the Quality of Next-Generation Sequencing in Clinical Microbiology and Public Health Laboratories.* Journal of Clinical Microbiology, 2016. **54**(12): p. 2857-2865.

12. Rehm, H.L., et al., *ACMG clinical laboratory standards for next-generation sequencing.* Genetics in medicine : official journal of the American College of Medical Genetics, 2013. **15**(9): p. 733-747.

13. Weiss, M.M., et al., *Best Practice Guidelines for the Use of Next-Generation Sequencing Applications in Genome Diagnostics: A National Collaborative Study of Dutch Genome Diagnostic Laboratories.* Human Mutation, 2013. **34**(10): p. 1313-1321.

14. Lowe, C.F., et al., *Implementation of Next-Generation Sequencing for Hepatitis B Virus Resistance Testing and Genotyping in a Clinical Microbiology Laboratory.* Journal of Clinical Microbiology, 2016. **54**(1): p. 127-133.

15. Deurenberg, R.H., et al., *Application of next generation sequencing in clinical microbiology and infection prevention.* Journal of Biotechnology, 2017. **243**: p. 16-24.

16. Sims, D.J., et al., *Plasmid-Based Materials as Multiplex Quality Controls and Calibrators for Clinical Next-Generation Sequencing Assays.* J Mol Diagn, 2016. **18**(3): p. 336-49.

17. Ong, F.S., et al., *Translational utility of next-generation sequencing.* Genomics, 2013. **102**(3): p. 137-9.

18. Moore, N.E., et al., *Metagenomic Analysis of Viruses in Feces from Unsolved Outbreaks of Gastroenteritis in Humans.* Journal of Clinical Microbiology, 2015. **53**(1): p. 15-21.

19. Scarpellini, E., et al., *The human gut microbiota and virome: Potential therapeutic implications.* Digestive and Liver Disease, 2015. **47**(12): p. 1007-1012.

20. Moustafa, A., et al., *The blood DNA virome in 8,000 humans.* PLOS Pathogens, 2017. **13**(3): p. e1006292.

21. Wylie, K.M., *The Virome of the Human Respiratory Tract.* Clin Chest Med, 2017. **38**(1): p. 11-19.

22. Hannigan, G.D., et al., *The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome.* MBio, 2015. **6**(5): p. e01578-15.

23. Oh, J., et al., *Biogeography and individuality shape function in the human skin metagenome.* Nature, 2014. **514**(7520): p. 59-64.

24. Foulongne, V., et al., *Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing.* PLoS One, 2012. **7**(6): p. e38499.

25. Wylie, K.M., et al., *Metagenomic analysis of double-stranded DNA viruses in healthy adults.* BMC Biology, 2014. **12**(1): p. 71.

26. Sloan, L.M., *Real-time PCR in clinical microbiology: verification, validation, and contamination control.* Clinical Microbiology Newsletter, 2007. **29**(12): p. 87-95.

27. Parker, J.Chen, J., *Application of next generation sequencing for the detection of human viral pathogens in clinical specimens.* Journal of Clinical Virology, 2017. **86**: p. 20-26.

28. Chesher, D., *Evaluating Assay Precision.* The Clinical Biochemist Reviews, 2008. **29**(Suppl 1): p. S23-S26.

29. Centers for Medicare and Medicaid Services. US Department of Health and Human Services. , C.-. *Part 493—Laboratory Requirements: Clinical Laboratory Improvement Amendments of 1988.* Fed Regist, 2004. **42**: p. 1443-1495.

30. Burd, E.M., *Validation of Laboratory-Developed Molecular Assays for Infectious Diseases.* Clinical Microbiology Reviews, 2010. **23**(3): p. 550-576

Chapter 1 Analytical sensitivity comparison between singleplex real-time PCR and a multiplex PCR platform for detecting respiratory viruses[1]

1.1 Abstract

Multiplex polymerase chain reaction (PCR) methods are attractive to clinical laboratories wanting to broaden their detection of respiratory viral pathogens in clinical specimens. However, multiplexed assays must be well optimized to retain or improve upon the analytic sensitivity of their singleplex counterparts. In this experiment, the lower limit of detection (LOD) of singleplex real-time PCR assays targeting respiratory viruses is compared to an equivalent panel on a multiplex PCR platform, the GenMark eSensor respiratory virus panel (RVP). LODs were measured for each singleplex real-time PCR assay and expressed as the lowest copy number detected 95–100% of the time, depending on the assay. The GenMark eSensor RVP LODs were obtained by converting the $TCID_{50}$/mL concentrations reported in the package insert to copies number equivalents using quantitative PCR (qPCR). LOD differences between the two methods ranged from 3 to 6,759 copies (0.46– 3.83 log difference) for all 12 assays compared. Assays targeting human parainfluenza 1 and 2 were most comparable (3 to 6 copies, < 1 log difference). Largest differences in LOD were demonstrated for assays targeting adenovirus group E, respiratory syncytial virus subtype A, and a generic assay for all influenza A viruses regardless of subtype (1,929 – 6,732 copies, 3.28–3.83 log difference). The multiplex PCR platform, the

[1] Parker, J., N. Fowler, M. L. Walmsley, T. Schmidt, J. Scharrer, J. Kowaleski, T. Grimes, S. Hoyos and J. Chen. *Analytical Sensitivity Comparison between Singleplex Real-Time PCR and a Multiplex PCR Platform for Detecting Respiratory Viruses*. PLOS ONE, 2015. **10**(11): e0143164. In this chapter, data was reanalyzed and some of the of the writing was modified to reflect these changes. The original publication can be found in Appendix A.

GenMark eSensor RVP, had better analytical sensitivity for detecting influenza A/H3 viruses, influenza B virus, and human rhinovirus with estimated detection improvement by 28 – 510 copies (1.45 to 2.70 logs). Broader detection of influenza A/H3 viruses was demonstrated by the GenMark eSensor RVP. The relationship between $TCID_{50}$/mL concentrations and the corresponding copy number related to various ATCC cultures is also reported.

1.2 Introduction

Multiplex PCR methods, those that target more than one pathogen in a single test, benefit diagnostics in a clinical laboratory due to their ability to detect and rule-out many related pathogens in the same amount of time. New and improved workflow designs make it possible for laboratories with varied molecular technical ability to implement multiplex PCR platforms.

The Respiratory Viral Panel (RVP) manufactured by GenMark Diagnostics, Inc. is a multiplex PCR panel that detects the amplification of various viral gene fragments electrochemically. Nucleic acids from targeted viral pathogens are amplified using a multiplex PCR reaction followed by denaturation of the double stranded molecules into single oligonucleotide strands using exonuclease. Once the amplicons are in a single-stranded state, they are hybridized to a complementary virus-specific signal probe tagged with ferrocene, a reducing agent. This hybridized molecule is then exposed to another sequence-specific probe which is bound to a solid phase, a gold electrode. Upon application of a low voltage current, the hybridized molecule bound to this solid phase brings the ferrocene in close proximity to the gold electrode where reversible electron transfer can occur and the resulting current can be measured. Viral pathogenic nucleic acid can be detected with confidence when measurements are ≥ 3 nanoamps (nA) on the GenMark XT-8 instrument. The GenMark eSensor RVP has been shown

to be highly comparable to other multiplex PCR platforms as well as singleplex real-time PCR in terms of diagnostic sensitivity and specificity [1, 2]. In this chapter, the primary interest is the analytical sensitivity of the PCR assays, or the minimum detectable concentration of the target. The GenMark eSensor RVP LODs as determined by the manufacturer are compared to singleplex real-time PCR assay LODs determined by our laboratory and expressed as the lowest copy number reliably detected 95–100% of the time.

Limit of detections for FDA-approved clinical assays, including those described in the GenMark eSensor RVP package insert, are commonly expressed as 50% tissue culture infectious dose per milliliter, or $TCID_{50}$/mL. Although this is a standard practice, other quantification methods such as real-time PCR are also reliable and may be able to more precisely describe quantities of viral particles with or without $TCID_{50}$/mL calculations as a reference [3-6]. Since the LODs for the GenMark RVP assays are expressed exclusively as $TCID_{50}$/mL concentrations, these values needed to be converted to copy number equivalents in order to meet our goals of comparing analytical sensitivity as lowest copy number. The LODs of each GenMark RVP assay were not re-established in our laboratory. Instead, manufacturer established $TCID_{50}$/mL values were converted to copy number using quantitative real-time PCR (qPCR). Performing this conversion also provided an opportunity to view the relationship between $TCID_{50}$/mL and copy number and relate this information to various virus-infected ATCC cell cultures.

The respiratory assays evaluated in this experiment target the following virus species: influenza A virus/H3, influenza A/H1N1, influenza B virus, human respiratory syncytial virus (RSV), human parainfluenza virus (serotypes 1, 2, and 3), human adenovirus, and human rhinovirus. The multiplex GenMark eSensor RVP assays were able to further distinguish human

adenoviruses as belonging to subgenera C or E and respiratory syncytial viruses as belonging to subgroup A or B, unlike the singleplex real-time PCR assays that were designed to detect human adenovirus and respiratory syncytial virus universally across all subgroups. A generic influenza A virus assay, one that targets a conserved region of all influenza A viruses regardless of subtype, was also evaluated.

1.3 Methods and Materials

1.3.1 Clinical specimens

Clinical specimens used in this study were de-identified. The University of Alaska Fairbanks Institutional Review Board (IRB) has determined that the proposed research qualifies for exemption from the requirements of 45 CFR 46 (Approval number: 667418–1).

1.3.2 Preparation of standard materials

Specific plasmids were created for each real-time PCR assay by ligating single copies of the diagnostic amplicon onto vectors (pCR 2.1 or pCR4, Invitrogen) and amplifying via TOPO cloning (Invitrogen). Transformant E.coli competent cells were extracted using a phenol/chloroform mixture and the presence of viral-specific inserts was verified by sequencing (Elim Biopharmaceuticals, Inc.).

Eleven plasmids were constructed and quantified using a gel electrophoresis method to exclude contaminating fragments of unwanted size as well as verify plasmid quality in terms of degradation. Due to their circular and supercoiled nature, the plasmids were linearized with restriction enzymes (either NcoI or SphI) to normalize their

movement through the gel matrix. An in-house standard was created from a plasmid of similar length (size ~4KB) as a comparator for plasmids of unknown quantity. A portion of the in-house standard was quantified using a fluorometer specific for DNA (Qubit 2.0) prior to running each gel, diluted, and added to wells as a standard curve. Using ImageJ software, a measurement tool with fixed area assigned values associated with pixel intensity of the fluorescing gel bands [7]. In addition to measuring the pixel intensity of each band, the intensity of a background image was also taken and subtracted from the initial measurement. The mass (in nanogram, ng) of the in-house standard was plotted against measured pixel intensities and a logarithmic trend line was used to interpolate the pixel intensity of the unknown plasmids and estimate their quantities in ng. Once mass was established for each gel band, the volume added to the gel was considered and a dilution factor was assigned based on the volume of stock plasmid represented in the restriction enzyme digests (total volume 25μL) to determine concentrations (ng/copy). The weight of each plasmid was calculated using Geneious (v.8.1.3) by taking into account the known sequence of the vector in addition to the Sanger-confirmed sequence of the insert. Final copy numbers concentrations (copies/μL) were calculated by dividing the plasmid weights (ng/copy) into the plasmid concentrations (ng/μL). Results of the quantification methods and downstream calculations are shown in Table 1.1 and Appendix 1.A.

1.3.3 Determination of singleplex real-time LOD

Stock plasmids were serially diluted to test a range of concentrations. Two identical dilution series were prepared, 1) diluted plasmids with nuclease-free water as

diluent and 2) diluted plasmids using pooled clinical specimen eluates (absent of the target) to simulate a clinical matrix. A narrow range of concentrations were chosen to identify the lowest potential copy number able to be detected repeatedly, but keep it above theoretical limitations of real time PCR, <3 copies (0.6 copies/µL when using 5µL reactions) [8]. Replicates (n=7) of each dilution were tested to determine the assay's LOD. Negative controls consisted of no template control replicates (NTC, n = 3) and diluent blank replicates, made up of water or patient eluate diluent (n = 7) to assess contamination. Positive reactions were defined as those amplification curves that produced cycle threshold (Ct) values at or below 40 cycles.

Primers and probes used in the laboratory-developed real-time PCR assays have been previously described [9, 10]. Influenza assays were performed using Invitrogen Superscript III reagents and all other assays were performed using Ambion AgPath ID reagents. For assays using the Invitrogen reagents, the following PCR thermal cycling profile was used; 50°C hold for 30 minutes, 95°C hold for 2 minutes, and 45 cycles of 95°C for 15 seconds then 55°C for 30 seconds. For assays using the Ambion reagents, the following PCR thermal cycling profile was used; 45°C hold for 10 minutes, 95°C for 10 minutes, and 45 cycles of 95°C for 15 seconds then 55°C for 1 minute. Reactions were tested using ABI 7500Dx thermal cyclers (Life Technologies).

The goal of these experiments was to determine the lowest copy number detectable by the assay 95-100% of the time, depending on the assay, to match the degree of positivity reported for the comparator GenMark assays. Most assays' LODs were set at 100% positivity, but three assays (influenza A/H1N1, RSVA, and

rhinovirus) were set lower and required further analysis to determine comparable LODs. This was done using probability units (probits) which are commonly used to show drug dose responses but can be used similarly for LOD determinations to evaluate assay performance (response) when provided various concentrations of DNA template (dose) [11]. Final LODs were expressed as copies per 5µL, the volume of template requred for each reaction (Table 1.2, Appendix 1.B).

1.3.4 Conversion of $TCID_{50}$/mL concentrations to copy number

Cell cultures with known $TCID_{50}$/mL quantities of target viruses (ATCC) were used to estimate the LOD for the GenMark RVP assay. Cultures were stored in liquid nitrogen until they were extracted using the easyMAG total nucleic acid automated extractor (Biomerieux). A total of 200µL of the $TCID_{50}$/mL culture was extracted and final eluate volumes were 60µL. Purified nucleic acid was stored at -80°C until tested by quantitative real-time PCR (qPCR).

Ten-fold dilutions of quantified plasmids containing inserts specific to each assay were tested in triplicate to create a standard curve (Appendix 1.C). All qPCR assays utilized a sequence-specific hydrolysis probe with the exception of the influenza A/H3 assay due to sequence incompatibilities with the ATCC strain being analyzed (see results). In this case, a SYBR Green assay (GoTaq, Promega) with new primers were designed to target this specific strain of influenza A/H3. Alongside the standard curve, dilutions of the isolated nucleic acid derived from the ATCC cultures were tested in triplicate at dilutions that would include reported GenMark eSensor RVP LOD $TCID_{50}$/mL values. As with the singleplex real-time PCR assays, reactions were tested

on ABI 7500Dx thermal cyclers (Life Technologies) and standard curves and associated unknown quantities were analyzed using ABI 7500 v2.3 software. The copy number equivalents for each GenMark eSensor assay's LOD is shown in Table 1.2. The relationship between copy number and $TCID_{50}$/mL for each ATCC culture tested is shown in Table 1.3.

1.4 Results

The analytical sensitivities of ten singleplex real-time PCR assays were compared to twelve multiplexed PCR GenMark eSensor RVP assays. The real-time PCR assays could not distinguish between different groups of adenovirus and respiratory syncytial viruses (RSV), while the GenMark eSensor RVP assay differentiates adenovirus group C vs. E and RSVA vs. RSVB, thus two more assays were evaluated for the GenMark eSensor RVP. Analytical sensitivity was expressed as lowest copy number for all assays.

Two of the twelve assays assessed, human parainfluenza 1 and 2, demonstrated similar performance between methods (<1 log copy number difference, Table 1.2). Three assays compared showed improved sensitivity using the multiplexed GenMark platform. These include the influenza A/H3, influenza B, and human rhinovirus assays which proved to have LODs exceeding that of the qPCR assay. Copy numbers reported for these assays are considered estimates since they were derived from continuing trend lines beyond the qPCR LOD. The GenMark eSensor human rhinovirus assay demonstrated the biggest difference in terms of improved detection when compared to its singleplex counterpart (estimated 510 ± 89 copies, 2.70 ± 0.1 log difference). Seven of the twelve assays compared showed greater sensitivity using the real-time singleplex assays. These include the adenovirus assays (C & E),

influenza A (generic), influenza A/H1N1, human parainfluenza 3, and RSV (A & B) with improvement ranging from 580 to 6,732 copies (2.75 to 3.83 log differences).

The Genmark eSensor RVP capable of distinguishing between different subgenera of adenoviruses (C vs. E) demonstrated less analytical sensitivity than the generic singleplex real-time PCR assay targeting all adenoviruses, differing by 1,033 ± 435 copies (2.99 ± 0.2 log difference), and 2,941 ± 379 copies (3.47 ± 0.6 log difference), respectively. The difference in sensitivity may be due to slight variations in the targeted priming region. The singleplex real-time PCR assays use primers designed to anneal highly conserved sequences within the hexon-coding region in order to target all adenoviruses, whereas the GenMark eSensor assays use subgenera-specific hexon primers to make the distinction between adenovirus group C and E possible. In terms of surveillance, differentiation of virus subtypes within a population may be important, regardless of lost sensitivity. Upper respiratory tract infections associated with adenovirus C viruses infect more than 80% of the population early in life [12]; however, infections with the adenovirus E serotype (type 4) can prove to be more severe and even fatal for people living in close quarters, such as military recruits [13]. In terms of surveillance, differentiation of virus subgenera within a population may be clinically useful, regardless of lost sensitivity.

Similarly, the singleplex real-time assay generically targeting RSV also demonstrated better sensitivity than the GenMark eSensor assays which are capable of distinguishing subtypes A and B (1,929 ± 154 copies, 3.28 ± 0.04 log difference and 580 ± 47 copies, 2.76 ± 0.04 log difference, respectively). Respiratory syncytial viruses in subtype A are thought to be more prevalent and virulent than those in subtype B [14]. Despite lower sensitivity, subtyping

respiratory syncytial virus may be beneficial when surveilling populations that experience high hospitalization rates associated with RSV, such as Native Americans living in southwest United States and Alaska [15].

Human parainfluenza 1 and 2 assays were highly comparable (6 ± 4 copies, 0.68 ± 0.3 log difference and 3 ± 2 copies, 0.46 ± 0.3 log difference, respectively). Human parainfluenza 3 assays demonstrated the largest difference in analytical sensitivity among the human parainfluenza serotypes, demonstrating a 2.86 ± 0.1 log improvement in detectability when using the singleplex real-time PCR assay (725 ± 101 copy difference).

Analytical sensitivity of assays targeting the current circulating strains of influenza A viruses in the human population, influenza A/H3 and influenza A/H1N1, demonstrated variable sensitivity between assays (76 ± 2 copies, 1.88 ± 0.01 log difference and 580 ± 149 copies, 2.75 ± 0.1 log difference, respectively), with the GenMark influenza A/H3 assay outperforming the singleplex assay in terms of sensitivity. Comparing the LOD between the influenza H3N2 assays proved to be the most challenging. When converting $TCID_{50}$/mL concentrations to copy numbers using qPCR, it was determined that this particular culture contained an uncommon virus, an Aichi strain (A/Aichi/2/35) circa 1968 (ATCC) and therefore could not be amplified using the singleplex real-time PCR assay, which is designed to detect current influenza A/H3N2 virus strains. However, it was repeatedly detected using the GenMark eSensor RVP. This finding suggests that the eSensor RVP is capable of detecting a broader range of influenza A/H3N2 strains while improving upon the analytic sensitivity of its singleplex real-time PCR counterpart.

The greatest difference measured between analytic sensitivities was seen with the generic influenza A assay showing a 3.83 ± 0.5 log difference in LOD (6,732 ± 843 copy difference) confirming better sensitivity with the singleplex assay. Because the LOD for the generic influenza A assay is much higher than the subtype assays (as described above) for the multiplex GenMark eSensor RVP, difficulty in result interpretation from specimens with low influenza A virus titers is likely (e.g. non-reactive influenza A (generic) reaction paired with a reactive H3N2 reaction). The performance of the generic influenza A assay is an important surveillance tool for tracking genetic changes among influenza A viruses. For instance, specimens demonstrating positivity for influenza A using this generic, highly conserved matrix-coding region may not subtype using the influenza A/H3 or influenza A/H1N1 assays, which may indicate that the virus is novel and worthy of alerting public health authorities. Therefore, it is imperitive for surveillance programs to use highly sensitive generic influenza A assays. In contrast, the influenza B assays were shown to be highly comparable between the singleplex and multiplex assays, with a difference of only 28 ± 3 copies (1.45 ± 0.1 log difference), showing slightly better sensitivity on the GenMark RVP assay.

The number of genome copies per $TCID_{50}$/mL value was highly variable ranging from 0.003 ± 0.0013 to 4,430 ± 56,329 (Table 1.3). LODs set at higher $TCID_{50}$/mL concentrations (10^3) corresponded to stock cultures with lower copy numbers (0.003 ± 0.0013 copies per $TCID_{50}$). LODs set at in the mid-range $TCID_{50}$/mL concentrations (10^1 to 10^{-1}) corresponded to stock cultures with variable copy numbers per $TCID_{50}$/mL (10 ± 1 to 5,686 ± 1,419 copies). LODs set at lower $TCID_{50}$/mL concentrations (10^{-2} to 10^{-3}) corresponded to stock cultures with higher copy numbers per $TCID_{50}$/ mL (391 ± 142 to 56,329 copies).

1.5 Conclusion

Multiplex PCR applications benefit diagnostics in a clinical laboratory due to their ability to detect and rule-out many related pathogens in a single reaction, reducing technician time by more than 3 hours for a panel of 10 viruses [1]. However, multiplex PCR platforms continue to struggle with loss in sensitivity. Analytic sensitivity, or the lowest possible concentration necessary to produce a reliable result, is an important parameter to consider when replacing singleplex realtime PCR assays with multiplex PCR platforms evolving from newer, more expensive technologies. This experiment aims at finding a method in which to compare LODs of various assays using copy number as the unit of expression.

Choosing a 2.5 log difference to express considerable loss in sensitivity, the multiplex PCR strategy in combination with the GenMark eSensor technology demonstrated a considerable loss in sensitivity for seven of the twelve assays assessed. Four of the assays were adenovirus groups C and E and respiratory syncytial virus subtypes A and B. Although sensitivity is reduced, further characterization of viruses in clinical specimens may be of greater clinical importance, especially when particular subtypes are known to be more virulent in the population as is the case with adenovirus serotype 4 (subgenera E) and respiratory syncytial virus subtype A in particular populations. The parainfluenza 3 assay also demonstrated loss in sensitivity, but showed it can be multiplexed without reducing the sensitivity of the parainfluenza serotypes 1 and 2 assays.

Two assays demonstrating considerable loss in sensitivity was the generic influenza A assay and, to a lesser extent, the influenza A/H1N1 assay. Clinical laboratories, especially those directly related to public health surveillance, may need to consider the significance of this

reduced sensitivity since it is commonly used to rule out novel influenza. Better analytic sensitivity was achieved using singleplex real-time PCR, which indicates that influenza A can be detected in clinical specimens even at low titers using this method. Specimens collected from patients that are suspected to have influenza infections that test negative on the GenMark eSensor RVP may need to be tested by more sensitive methods to rule out cases of novel influenza.

Expressing LOD in units that can be comparable across methodologies can prove to be difficult experimentally. $TCID_{50}$/mL measurements can vary depending on how these cultures are handled in the laboratory in regards to preserving the concentration of infectious virus particles for purposes of experimentation and quantity comparisons. Molecular detection strategies used in clinical laboratories are non-discriminating when identifying nucleic acid. PCR methodologies used to detect viral targets in clinical specimens cannot determine the viability of the virus and, therefore, every detection may not point to a causative agent of disease. Other complicating factors to consider when interpreting PCR results are that patients can be asymptomatic carriers or may be exhibiting evidence of a past infections. Viral copy numbers provide an estimate of the number of virus particles in a given volume, but in our experiment, they did not correlate well with the number of infectious particles. To test the analytical sensitivity of a PCR-based methodology, it is important to understand that the intent of the assay is to detect any genome copy targeted by the designed primers, whether these be from infectious or non-infectious virus particles.

1.6 Acknowledgments

This research was supported by the Alaska Department of Health and Social Services, Division of Public Health, Section of laboratories. Much of the plasmid development and sequencing was supported in part by the University of Alaska Fairbanks. The ATCC cultures were purchased by GenMark Diagnostics, Incorporated in an effort to be consistent with the particular strains used in the FDA validation testing. We would like to thank the staff at the Alaska State Virology Laboratory for all of their help with carrying out testing for this project.

Table 1.1: Plasmid concentrations and copy number determination

Virus target insertion	Vector	ImageJ concentration (ng/µL)	Weight per plasmid copy (ng)	Copies/µL	Stock solution concentration (ng/µL)	Stock copies/µL
Adenovirus	pCR2.1	48.9	4.17×10^{-9}	1.17×10^{10}	7.5	1.80×10^{9}
Influenza A (generic)	pCR2.1	23.5	4.14×10^{-9}	5.68×10^{9}	10	2.41×10^{9}
Influenza A/H3	pCR4	9.4	4.37×10^{-9}	3.02×10^{9}	7.5	1.72×10^{9}
Influenza A/H1N1	pCR2.1	10.8	4.15×10^{-9}	2.60×10^{9}	10	2.41×10^{9}
Influenza B	pCR4	38.1	4.17×10^{-9}	9.14×10^{9}	25	6.00×10^{9}
Parainfluenza 1	pCR4	72.0	4.32×10^{-9}	1.67×10^{10}	10	2.31×10^{9}
Parainfluenza 2	pCR4	103.6	4.15×10^{-9}	2.50×10^{10}	5	1.20×10^{9}
Parainfluenza 3	pCR4	31.6	4.19×10^{-9}	7.54×10^{9}	5	1.19×10^{9}
RSVA	pCR4	48.5	4.28×10^{-9}	1.13×10^{10}	10	2.34×10^{9}
RSVB	pCR4	9.3	4.28×10^{-9}	$2.17 \text{ X } 10^{9}$	5	1.17×10^{9}
Rhinovirus	pCR4	44.4	4.27×10^{-9}	1.04×10^{10}	30	7.02×10^{9}

RSV = respiratory syncytial virus. Weight per copy was calculated using Geneious (v.8.1.3) which considers the exact sequence of the plasmid, including the specific virus target insertion. See Appendix 1.A describing calculations. This table has been updated since the original publication which can be found in Appendix A.

Table 1.2: Limit of detection (LOD) comparison summary

Assay	% pos	Singleplex LOD in clinical background (copies)	Singleplex LOD in water background (copies)	GenMark RVP LOD (copy number equivalent of $TCID_{50}$/mL)	Copy number difference	Log Difference
Similar assay performance (< 1 log difference)						
Parainfluenza 1	100%	5	5	11 ± 4	6 ± 4	0.68 ± 0.3
Parainfluenza 2	100%	27	11	28 ± 4	3 ± 2	0.46 ± 0.3
GenMark RVP assays with better sensitivity						
Influenza A/H3	100%	80	40	4 ± 2 (est.)	76 ± 2	1.88 ± 0.01
Influenza B	100%	33	660	5 ± 3 (est.)	28 ± 3	1.45 ± 0.1
Rhinovirus	97.5%	517	263	7 ± 89 (est.)	510 ± 89	2.70 ± 0.1
Singleplex real-time PCR assays with better sensitivity						
Adenovirus C	100%	8	20	1,041 ± 435	1,033 ± 435	2.99 ± 0.2
Adenovirus E	100%	8	20	2,948 ± 379	2,941 ± 379	3.47 ± 0.6
Influenza A (generic)	100%	27	106	6,759 ± 843	6,732 ± 843	3.83 ± 0.5
Influenza A/H1N1	95%	34	34	597 ± 149	580 ± 149	2.75 ± 0.1
Parainfluenza 3	100%	11	5	736 ± 101	725 ± 101	2.86 ± 0.1
RSVA	97.5%	27	10	1,952 ± 188	1,929 ± 154	3.28 ± 0.04
RSVB	100%	11	26	591 ± 47	580 ± 47	2.76 ± 0.04

Lowest copy numbers are expressed as the nearest whole number of copies in 5µL of eluate used in each reaction. Adenovirus assays were not differentiated with the singleplex real-time PCR assay. The $TCID_{50}$/mL concentration for influenza A/H3, influenza B, and rhinovirus exceeded the detection limit of the qPCR assay and are considered estimates (est.) of continued trend lines beyond the qPCR LOD. Copy number differences were calculated using the singleplex LODs in clinical background in comparison to the average copy number equivalent of the GenMark RVP LODs. See Appendix 1.B for singleplex assay calculation details and Appendix 1.C describing TCID conversions to copy numbers. This table has been updated since the original publication which can be found in Appendix A.

Table 1.3: Relationship between $TCID_{50}$/mL concentrations and copy number

ATCC Culture		LOD for GenMark eSensor assays ($TCID_{50}$/mL)	Copy equivalent at LOD concentration	Copies per $TCID_{50}$
VR-1	Adenovirus Type 1 (C)	8.89×10^{1}	1,041 ± 435	12 ± 5
VR-1572	Adenovirus Type 4 (E)	1.58×10^{1}	2,948 ± 379	187 ± 24
VR-547	Influenza A/H3 (Aichi)	1.58×10^{3}	4 ± 2	0.003 ± 0.0013
VR-1736	Influenza A/H1N1	1.05×10^{-1}	597 ± 149	5,686 ± 1,419
VR-101	Influenza B	3.16×10^{-1}	5 ± 3	16 ± 9
VR-94	Human Parainfluenza Virus Type 1 (C35)	2.81×10^{-2}	11 ± 4	391 ± 142
VR-92	Human Parainfluenza Virus Type 2 (Greer)	2.81×10^{0}	28 ± 4	10 ± 1
VR-93	Human Parainfluenza Virus Type 3 (C243)	2.81×10^{1}	736 ± 101	26 ± 4
VR-1540	Respiratory Syncytial Virus (A2)	2.81×10^{0}	1,952 ± 188	695 ± 67
VR-955	Respiratory Syncytial Virus (B9320)	1.58×10^{0}	591 ± 47	374 ± 30
VR-483	Rhinovirus 3 FEB	1.58×10^{-3}	7 ± 89	4,430 ± 56,329

Using the equivalent copy number determination at the LOD $TCID_{50}$/mL concentration for each assay, the number of copies per $TCID_{50}$/mL designation is estimated with standard error. This table has been updated since the original publication which can be found in Appendix A.

1.7 References

1. Pierce, V.M. and R.L. Hodinka, *Comparison of the GenMark Diagnostics eSensor Respiratory Viral Panel to Real-Time PCR for Detection of Respiratory Viruses in Children.* Journal of Clinical Microbiology, 2012. **50**(11): p. 3458-3465.

2. Popowitch, E.B., S.S. O'Neill, and M.B. Miller, *Comparison of the Biofire FilmArray RP, Genmark eSensor RVP, Luminex xTAG RVPv1, and Luminex xTAG RVP Fast Multiplex Assays for Detection of Respiratory Viruses.* Journal of Clinical Microbiology, 2013. **51**(5): p. 1528-1533.

3. Jonsson, N., M. Gullberg, and A.M. Lindberg, *Real-time polymerase chain reaction as a rapid and efficient alternative to estimation of picornavirus titers by tissue culture infectious dose 50% or plaque forming units.* Microbiology and Immunology, 2009. **53**(3): p. 149-154.

4. Gustafsson, R.K.L., E.E. Engdahl, and A. Fogdell-Hahn, *Development and validation of a Q-PCR based TCID(50) method for human herpesvirus 6.* Virology Journal, 2012. **9**: p. 311-311.

5. Kallesh, D., et al., *Quantitative PCR: A quality control assay for estimation of viable virus content in live attenuated goat pox vaccine.* Indian journal of experimental biology, 2009. **47**(11): p. 911.

6. Iwami, S., et al., *Quantification system for the viral dynamics of a highly pathogenic simian/human immunodeficiency virus based on an in vitro experiment and a mathematical model.* Retrovirology, 2012. **9**(1): p. 18.

7. Rasband, W.S., *ImageJ,* U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2014.

8. Bustin, S.A., et al., *The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments.* Clinical Chemistry, 2009. **55**(4): p. 611-622.

9. Weinberg, G.A., et al., *Field evaluation of TaqMan Array Card (TAC) for the simultaneous detection of multiple respiratory viruses in children with acute respiratory infection.* J Clin Virol, 2013. **57**(3): p. 254-60.

10. Wangchuk, S., et al., *Influenza surveillance from November 2008 to 2011; including pandemic influenza A(H1N1)pdm09 in Bhutan.* Influenza Other Respir Viruses, 2013. **7**(3): p. 426-30.

11. Sloan, L.M., *Real-time PCR in clinical microbiology: verification, validation, and contamination control.* Clinical Microbiology Newsletter, 2007. **29**(12): p. 87-95.

12. Garnett, C.T., et al., *Prevalence and Quantitation of Species C Adenovirus DNA in Human Mucosal Lymphocytes.* Journal of Virology, 2002. **76**(21): p. 10608-10616.

13. Robert, N.P., et al., *Adenovirus-associated Deaths in US Military during Postvaccination Period, 1999–2010.* Emerging Infectious Disease journal, 2012. **18**(3): p. 507.

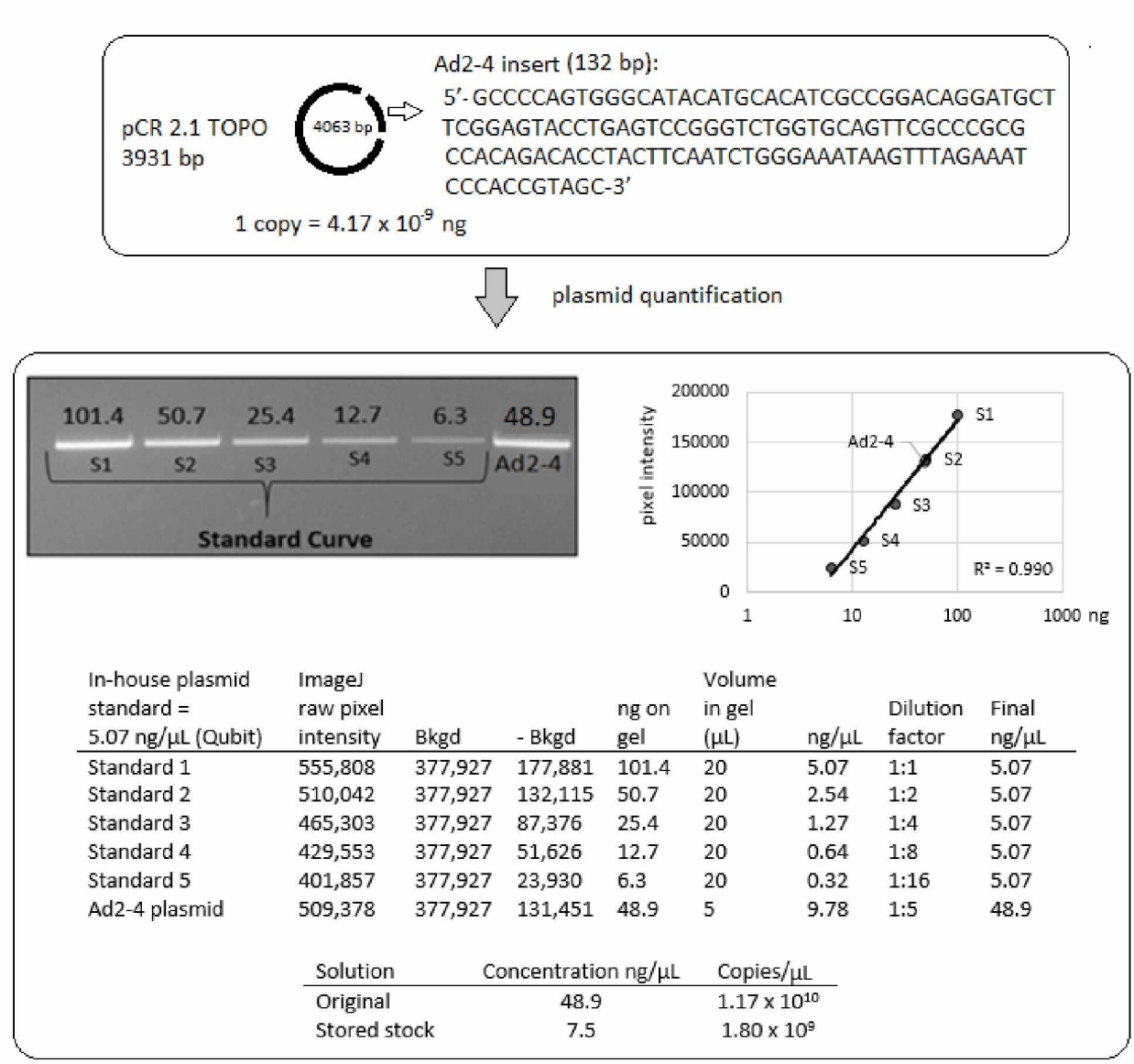
14. Walsh, E.E., et al., *Severity of Respiratory Syncytial Virus Infection Is Related to Virus Strain.* Journal of Infectious Diseases, 1997. **175**(4): p. 814-820.

15. Holman, R.C., et al., *Respiratory syncytial virus hospitalizations among American Indian and Alaska Native infants and the general United States infant population.* Pediatrics, 2004. **114**(4): p. e437-44.

Appendix 1.A

Plasmid construction and quantification details

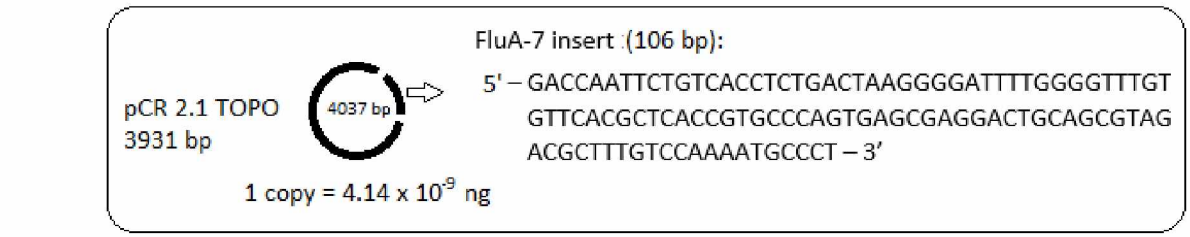
This appendix describes the laboratory details of constructing eleven plasmids containing inserts from respiratory viruses and how they were quantified. Once original concentrations were assigned for each plasmid, stock solutions were prepared for use in the LOD experiments (Appendix 1.B and 1.C). In the following figures, gel images show mass in nanograms over each gel band for the standard curve and the amount estimated for each individual respiratory virus plasmid. Tables below each gel image list pixel intensity measurements for each band alongside a logarithmic plot used to determine unknown plasmid quantities based on the pixel intensity of gel band images. Final concentrations and corresponding copy numbers are listed at the bottom.



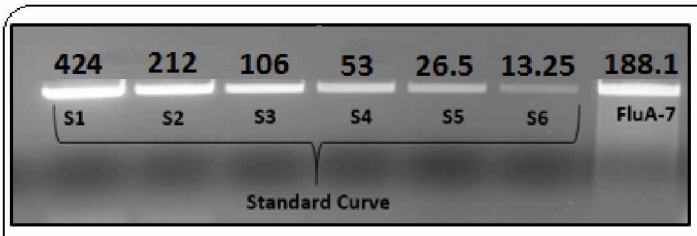
In-house plasmid standard = 5.07 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	555,808	377,927	177,881	101.4	20	5.07	1:1	5.07
Standard 2	510,042	377,927	132,115	50.7	20	2.54	1:2	5.07
Standard 3	465,303	377,927	87,376	25.4	20	1.27	1:4	5.07
Standard 4	429,553	377,927	51,626	12.7	20	0.64	1:8	5.07
Standard 5	401,857	377,927	23,930	6.3	20	0.32	1:16	5.07
Ad2-4 plasmid	509,378	377,927	131,451	48.9	5	9.78	1:5	48.9

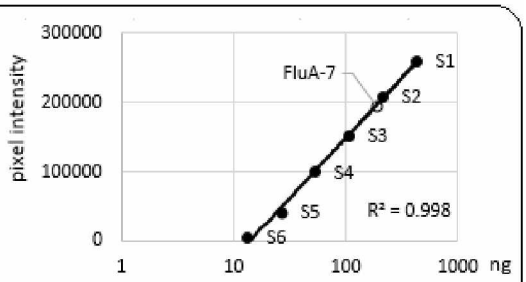
Solution	Concentration ng/µL	Copies/µL
Original	48.9	1.17×10^{10}
Stored stock	7.5	1.80×10^{9}

Figure 1.A-1: Adenovirus (Ad2-4) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



plasmid quantification

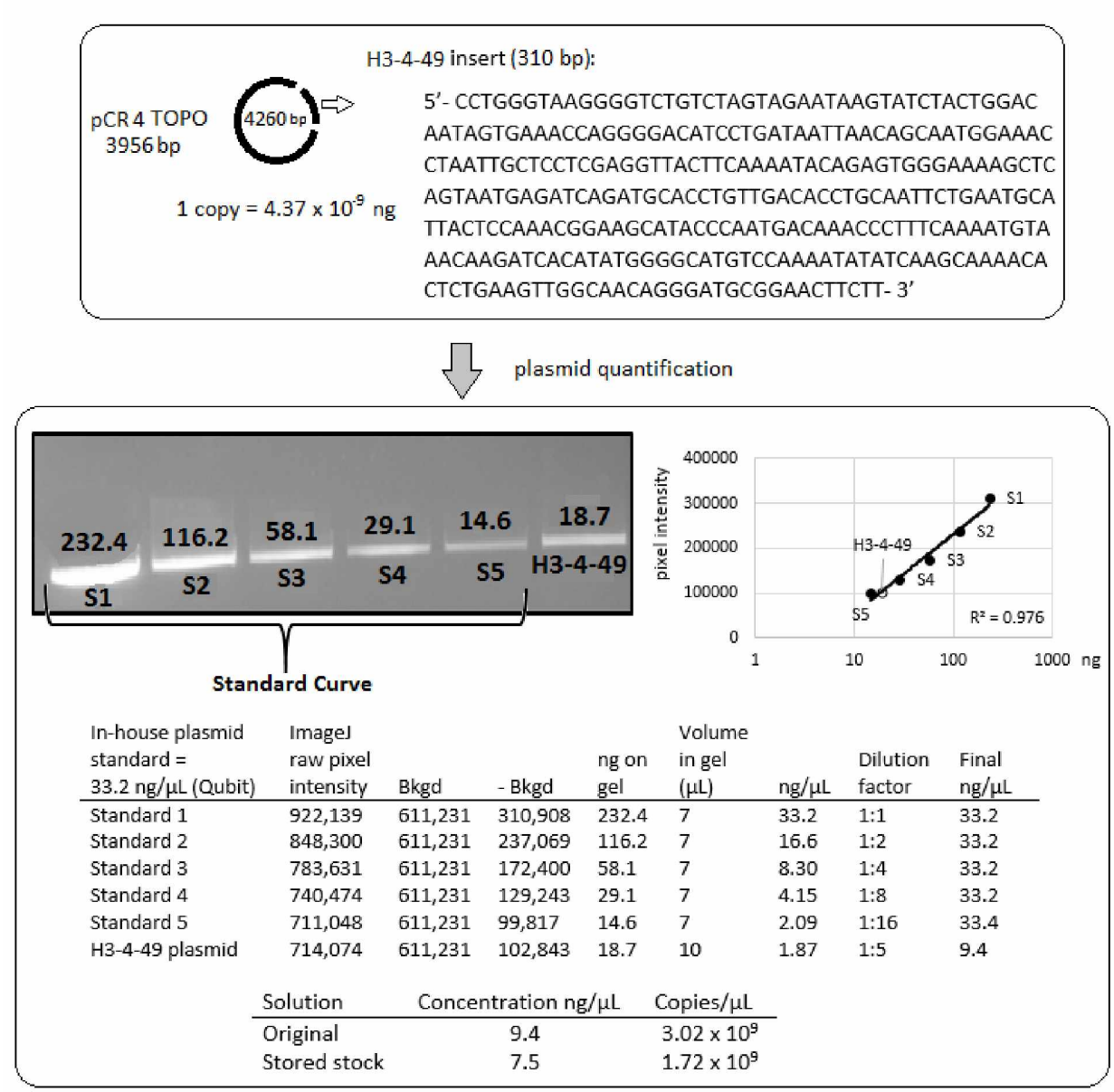




In-house plasmid standard = 42.4 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	677,925	418,709	259,216	424	20	21.2	1:2	42.4
Standard 2	625,689	418,709	206,980	212	20	10.6	1:4	42.4
Standard 3	570,243	418,709	151,534	106	20	5.3	1:8	42.4
Standard 4	517,692	418,709	98,983	53	20	2.65	1:16	42.4
Standard 5	460,425	418,709	41,716	26.5	20	13.25	1:32	42.4
Standard 6	422,644	418,709	3,935	13.25	20	0.663	1:64	42.4
FluA-7 plasmid	615,124	418,709	196,415	188.1	20	9.4	1:2.5	23.5

Solution	Concentration ng/µL	Copies/µL
Original	23.5	5.68×10^9
Stored stock	10.0	2.41×10^9

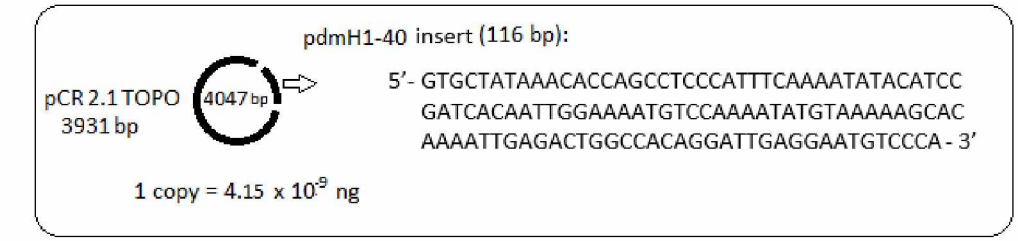
Figure 1.A-2: Influenza A (FluA-7) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



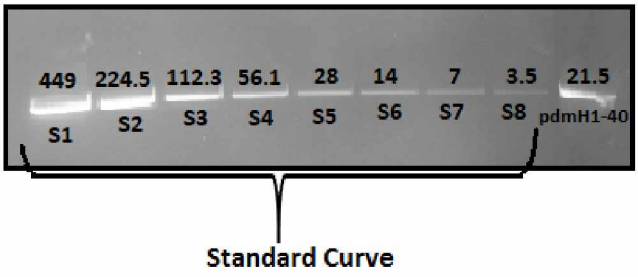
In-house plasmid standard = 33.2 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	922,139	611,231	310,908	232.4	7	33.2	1:1	33.2
Standard 2	848,300	611,231	237,069	116.2	7	16.6	1:2	33.2
Standard 3	783,631	611,231	172,400	58.1	7	8.30	1:4	33.2
Standard 4	740,474	611,231	129,243	29.1	7	4.15	1:8	33.2
Standard 5	711,048	611,231	99,817	14.6	7	2.09	1:16	33.4
H3-4-49 plasmid	714,074	611,231	102,843	18.7	10	1.87	1:5	9.4

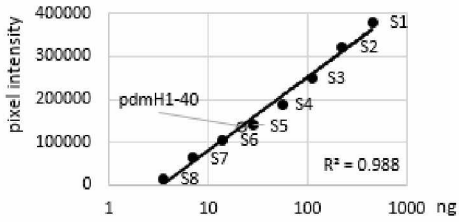
Solution	Concentration ng/µL	Copies/µL
Original	9.4	3.02×10^9
Stored stock	7.5	1.72×10^9

Figure 1.A-3: Influenza A/H3 (H3-4-49) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



plasmid quantification



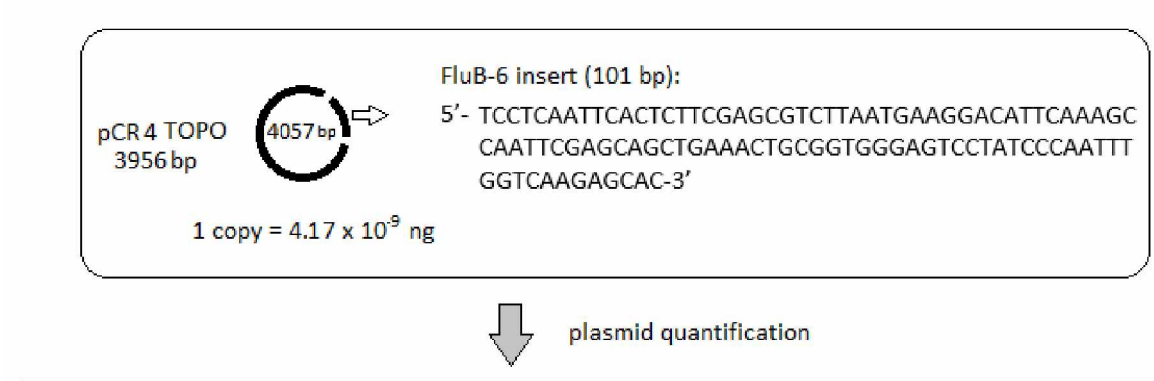


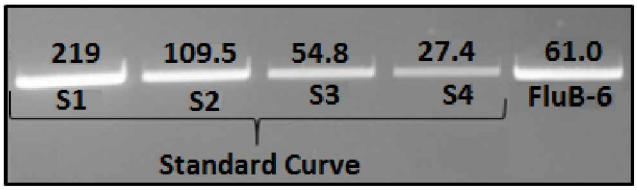
Standard Curve

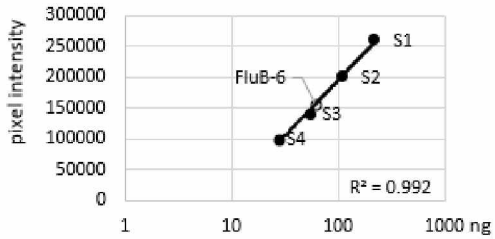
In-house plasmid standard = 44.8 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution Factor	Final ng/µL
Standard 1	851,058	471,455	379,603	449	10	44.9	1:1	44.9
Standard 2	79,3063	471,455	321,608	224.5	10	22.4	1:2	44.8
Standard 3	720,878	471,455	249,423	112.3	10	11.2	1:4	44.8
Standard 4	660,847	471,455	189,392	56.1	10	5.6	1:8	44.8
Standard 5	610,328	471,455	138,873	28	10	2.8	1:16	44.8
Standard 6	577,068	471,455	105,613	14	10	1.4	1:32	44.8
Standard 7	534,639	471,455	63,184	7	10	0.7	1:64	44.8
Standard 8	486,240	471,455	14,785	3.5	10	0.35	1:128	44.8
pdmH1-40 plasmid	608,795	471,455	137,340	21.5	5	4.3	1:2.5	10.8

Solution	Concentration ng/µL	Copies/µL
Original	10.8	2.60×10^9
Stored stock	10.0	2.41×10^9

Figure 1.A-4: Influenza A/H1N1 (pdmH1-40) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



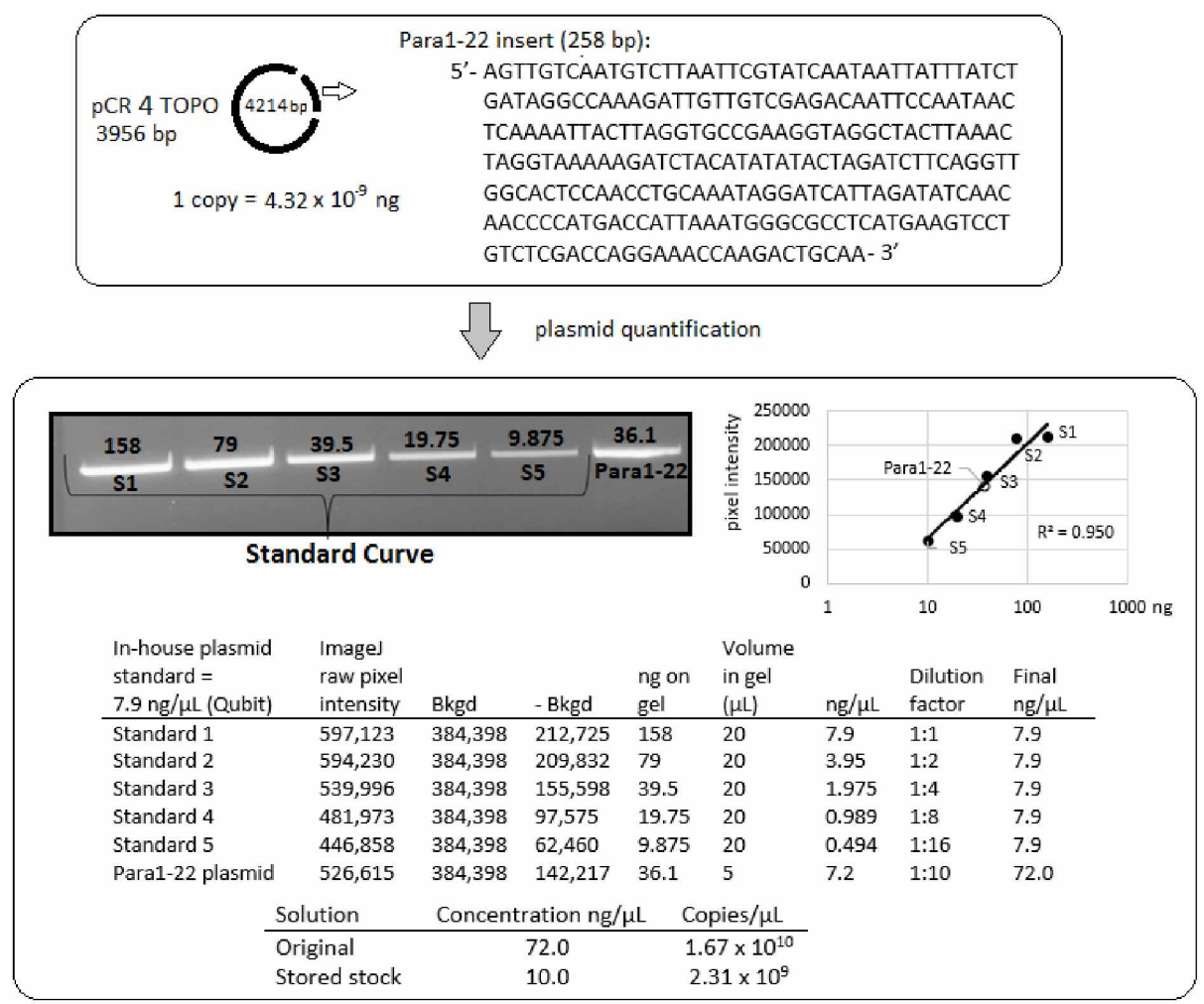




In-house plasmid standard = 21.9 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	751,435	490,436	260,999	219	10	21.9	1:1	21.9
Standard 2	691,793	490,436	201,357	109.5	10	10.95	1:2	21.9
Standard 3	629,799	490,436	139,363	54.8	10	5.48	1:4	21.9
Standard 4	590,980	490,436	100,544	27.4	10	2.74	1:8	21.9
FluB-6 plasmid	646,925	490,436	156,489	61.0	4	15.25	1:2.5	38.1

Solution	Concentration ng/µL	Copies/µL
Original	38.1	9.14×10^9
Stored stock	25.0	6.00×10^9

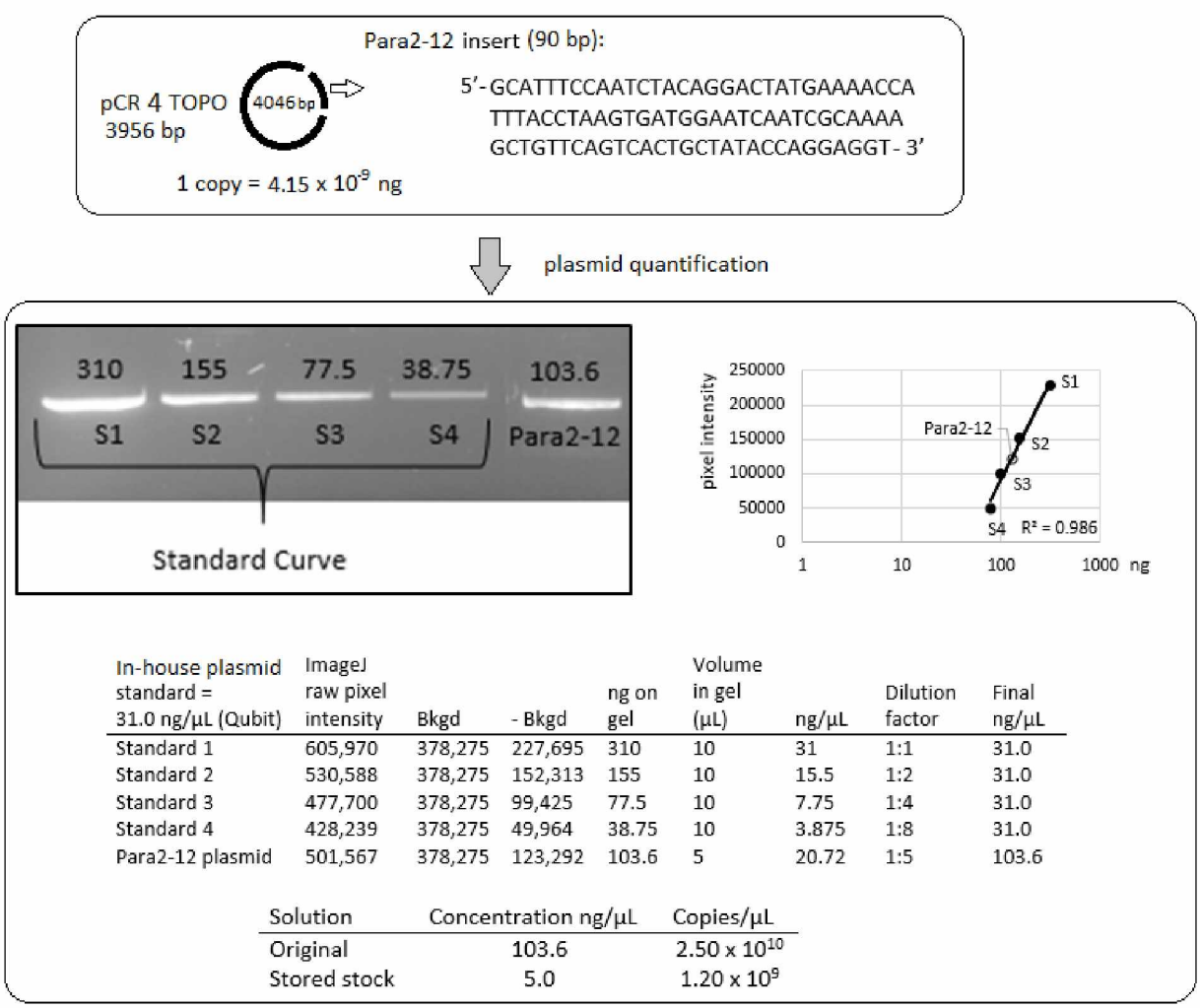
Figure 1.A-5: Influenza B (FluB-6) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



In-house plasmid standard = 7.9 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	597,123	384,398	212,725	158	20	7.9	1:1	7.9
Standard 2	594,230	384,398	209,832	79	20	3.95	1:2	7.9
Standard 3	539,996	384,398	155,598	39.5	20	1.975	1:4	7.9
Standard 4	481,973	384,398	97,575	19.75	20	0.989	1:8	7.9
Standard 5	446,858	384,398	62,460	9.875	20	0.494	1:16	7.9
Para1-22 plasmid	526,615	384,398	142,217	36.1	5	7.2	1:10	72.0

Solution	Concentration ng/µL	Copies/µL
Original	72.0	1.67×10^{10}
Stored stock	10.0	2.31×10^{9}

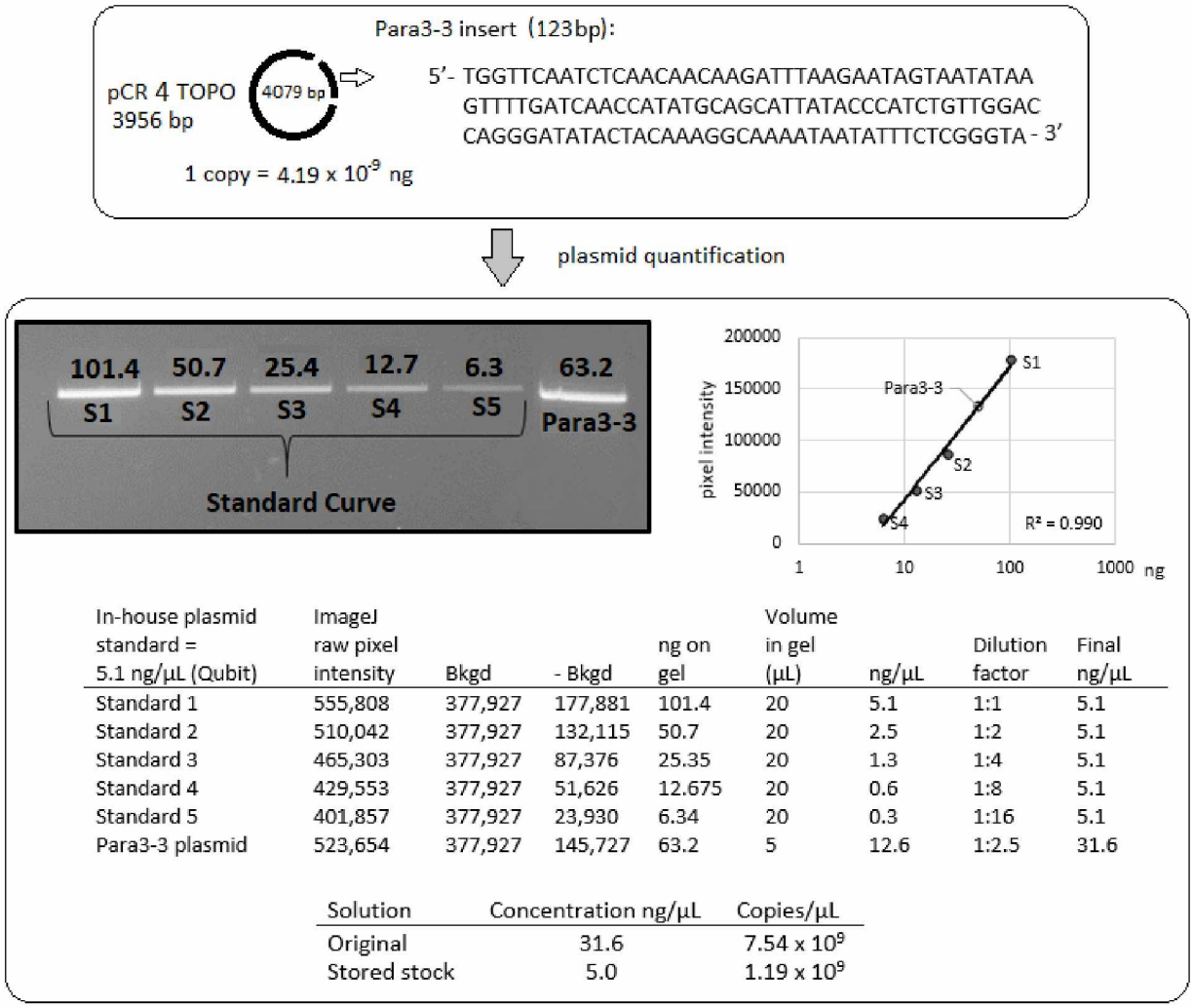
Figure 1.A-6: Parainfluenza 1 (Para1-22) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



In-house plasmid standard = 31.0 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	605,970	378,275	227,695	310	10	31	1:1	31.0
Standard 2	530,588	378,275	152,313	155	10	15.5	1:2	31.0
Standard 3	477,700	378,275	99,425	77.5	10	7.75	1:4	31.0
Standard 4	428,239	378,275	49,964	38.75	10	3.875	1:8	31.0
Para2-12 plasmid	501,567	378,275	123,292	103.6	5	20.72	1:5	103.6

Solution	Concentration ng/µL	Copies/µL
Original	103.6	2.50×10^{10}
Stored stock	5.0	1.20×10^{9}

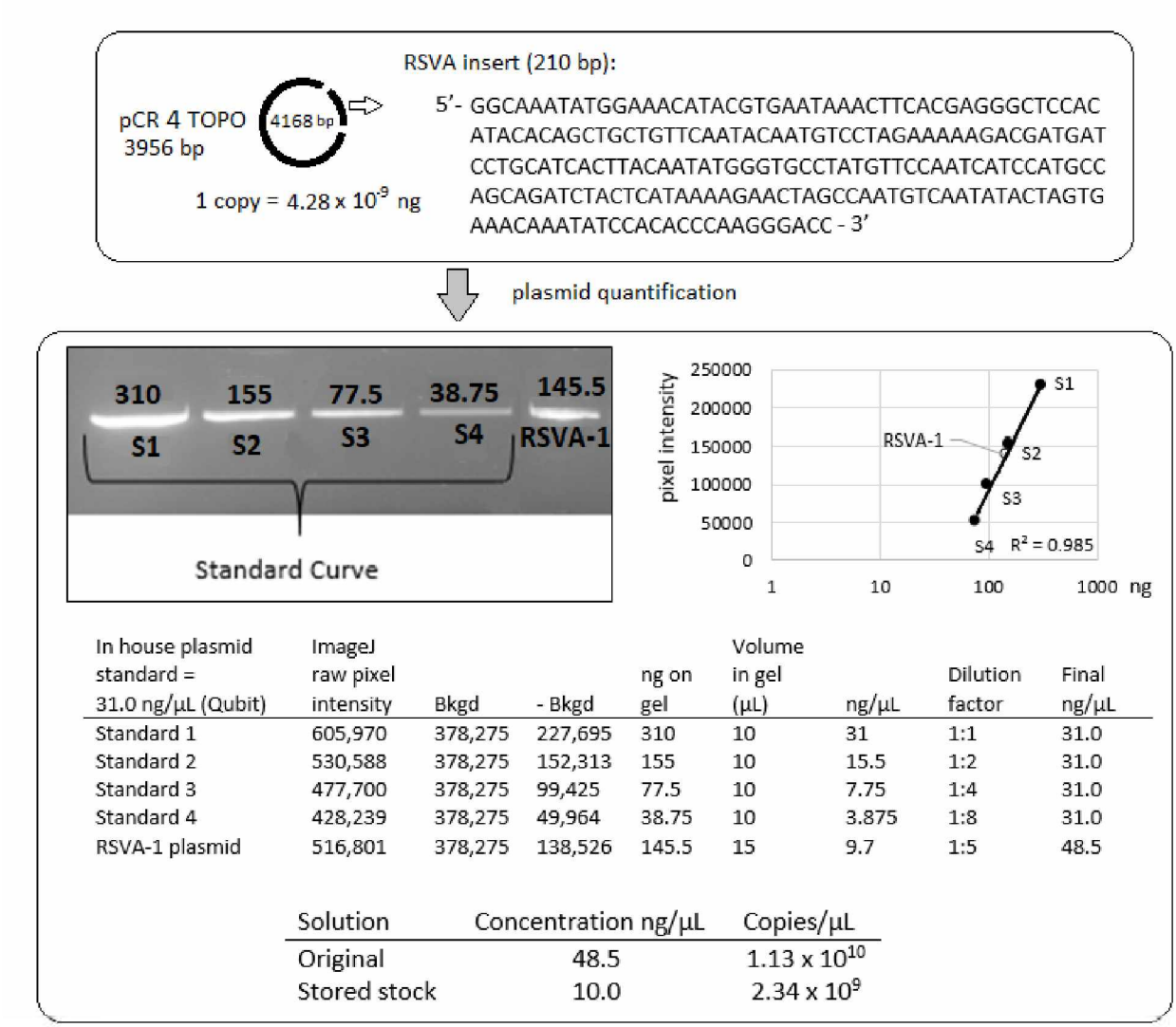
Figure 1.A-7: Parainfluenza 2 (Para2-12) plasmid details. Plasmid was linearized using restriction enzyme SphI.



In-house plasmid standard = 5.1 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	555,808	377,927	177,881	101.4	20	5.1	1:1	5.1
Standard 2	510,042	377,927	132,115	50.7	20	2.5	1:2	5.1
Standard 3	465,303	377,927	87,376	25.35	20	1.3	1:4	5.1
Standard 4	429,553	377,927	51,626	12.675	20	0.6	1:8	5.1
Standard 5	401,857	377,927	23,930	6.34	20	0.3	1:16	5.1
Para3-3 plasmid	523,654	377,927	145,727	63.2	5	12.6	1:2.5	31.6

Solution	Concentration ng/µL	Copies/µL
Original	31.6	7.54×10^9
Stored stock	5.0	1.19×10^9

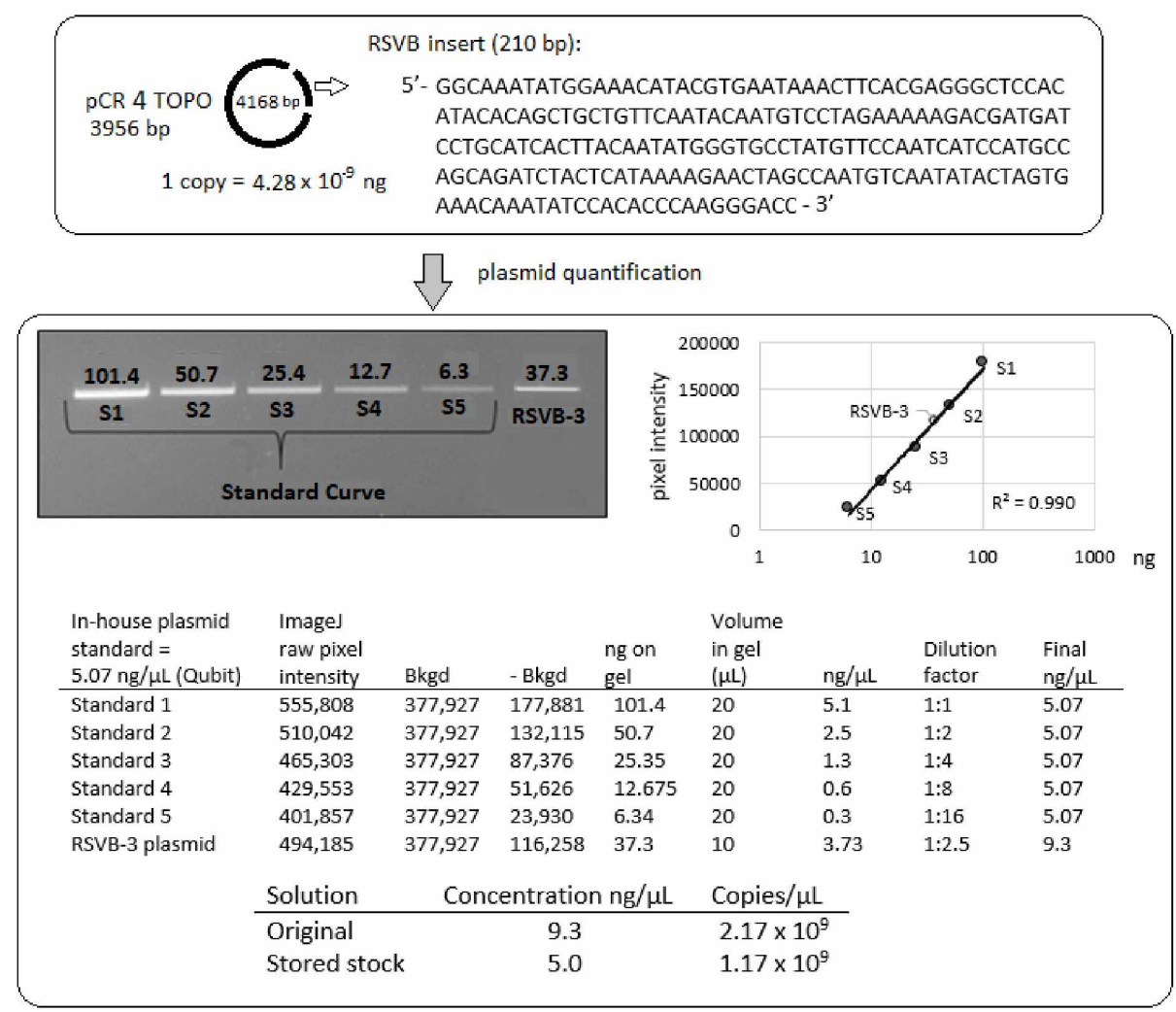
Figure 1.A-8: Parainfluenza 3 (Para3-3) plasmid details. Plasmid was linearized using restriction enzyme NcoI.



In house plasmid standard = 31.0 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	605,970	378,275	227,695	310	10	31	1:1	31.0
Standard 2	530,588	378,275	152,313	155	10	15.5	1:2	31.0
Standard 3	477,700	378,275	99,425	77.5	10	7.75	1:4	31.0
Standard 4	428,239	378,275	49,964	38.75	10	3.875	1:8	31.0
RSVA-1 plasmid	516,801	378,275	138,526	145.5	15	9.7	1:5	48.5

Solution	Concentration ng/µL	Copies/µL
Original	48.5	1.13×10^{10}
Stored stock	10.0	2.34×10^{9}

Figure 1.A-9: Respiratory syncytial virus A (RSVA-1) plasmid details. Plasmid was linearized using restriction enzyme SphI.



In-house plasmid standard = 5.07 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/µL
Standard 1	555,808	377,927	177,881	101.4	20	5.1	1:1	5.07
Standard 2	510,042	377,927	132,115	50.7	20	2.5	1:2	5.07
Standard 3	465,303	377,927	87,376	25.35	20	1.3	1:4	5.07
Standard 4	429,553	377,927	51,626	12.675	20	0.6	1:8	5.07
Standard 5	401,857	377,927	23,930	6.34	20	0.3	1:16	5.07
RSVB-3 plasmid	494,185	377,927	116,258	37.3	10	3.73	1:2.5	9.3

Solution	Concentration ng/µL	Copies/µL
Original	9.3	2.17×10^9
Stored stock	5.0	1.17×10^9

Figure 1.A-10: Respiratory syncytial virus B (RSVB-3) plasmid details. Plasmid was linearized using restriction enzyme NcoI.

In-house plasmid standard = 33.2 ng/µL (Qubit)	ImageJ raw pixel intensity	Bkgd	- Bkgd	ng on gel	Volume in gel (µL)	ng/µL	Dilution factor	Final ng/ µL
Standard 1	922,139	611,231	310,908	232.4	7	33.2	1:1	33.2
Standard 2	848,300	611,231	237,069	116.2	7	16.6	1:2	33.2
Standard 3	783,631	611,231	172,400	58.1	7	8.3	1:4	33.2
Standard 4	740,474	611,231	129,243	29.1	7	4.2	1:8	33.2
Standard 5	711,048	611,231	99,817	14.7	7	2.1	1:16	33.2
Rhino-2 plasmid	763,106	611,231	151,875	35.5	10	3.55	1:12.5	44.4

Solution	Concentration ng/µL	Copies/µL
Original	44.4	1.04×10^{10}
Stored stock	30.0	7.02×10^{9}

Figure 1.A-11: Rhinovirus (Rhino-2) plasmid details. Plasmid was linearized using restriction enzyme NcoI.

Appendix 1.B

Establishing limits of detection for singleplex real-time PCR assays

Plasmids described in Appendix 1.A were used to determine the limit of detection (LOD) of ten singleplex real-time PCR assays targeting human respiratory viruses. Two separate RSV plasmids were created (RSVA and RSVB) and tested on the same real-time PCR assay. This appendix describes the laboratory details of each LOD experiment. In the following figures, positive reactions are depicted as black circles and negative reactions are shown as empty circles. The percentage of reactions that are positive for each dilution is shown below each column. Three assays underwent probit analysis for LODs set at positivity levels less <100%, influenza A/H1N1 (95%), RSVA (97.5%), and human rhinovirus (97.5%).

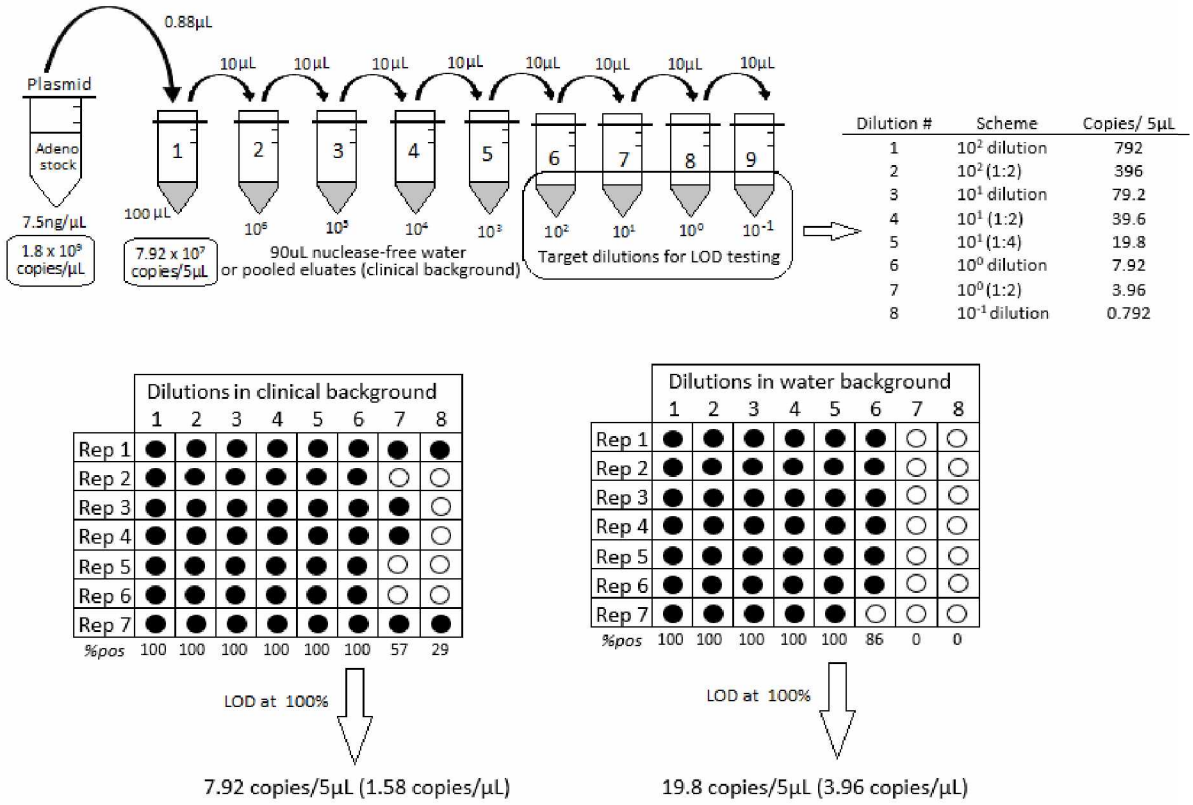


Figure 1.B-1: Adenovirus singleplex real-time PCR assay limit of detection determination

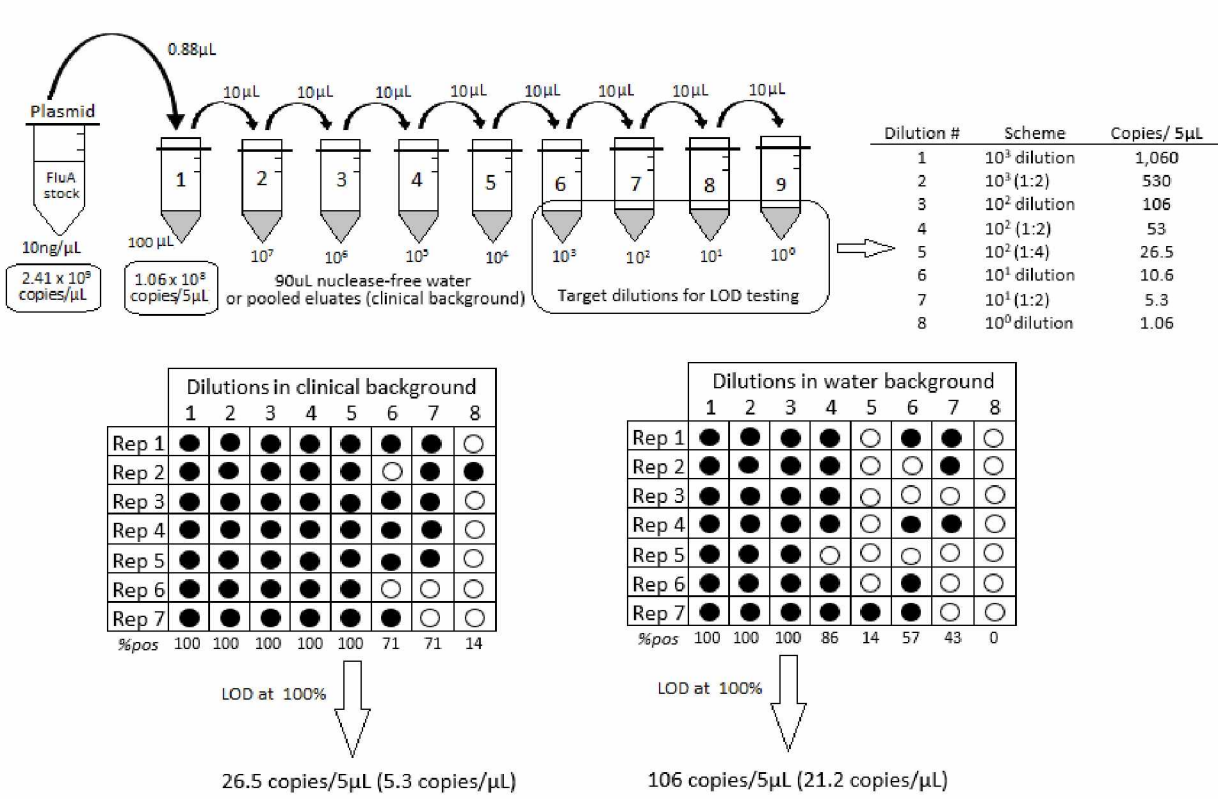


Figure 1.B-2: Influenza A (generic) singleplex real-time PCR assay limit of detection determination

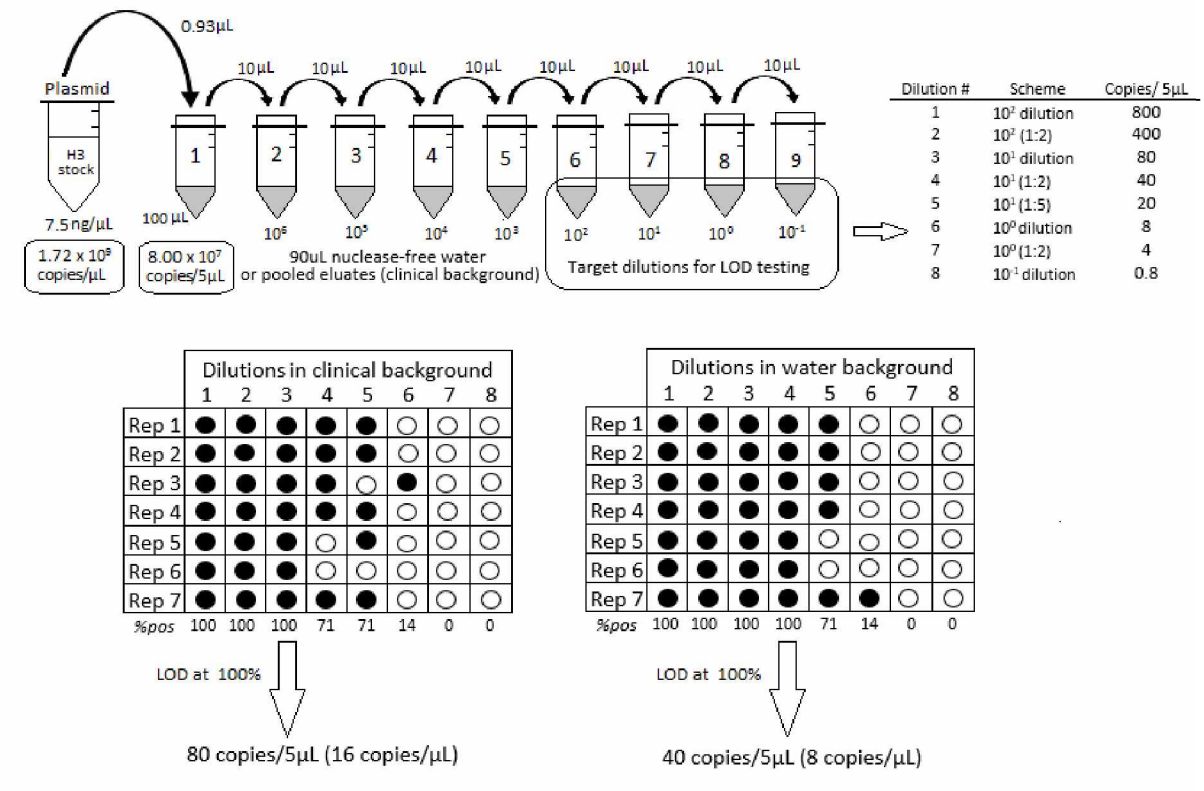


Figure 1.B-3: Influenza A/H3 singleplex real-time PCR assay limit of detection determination

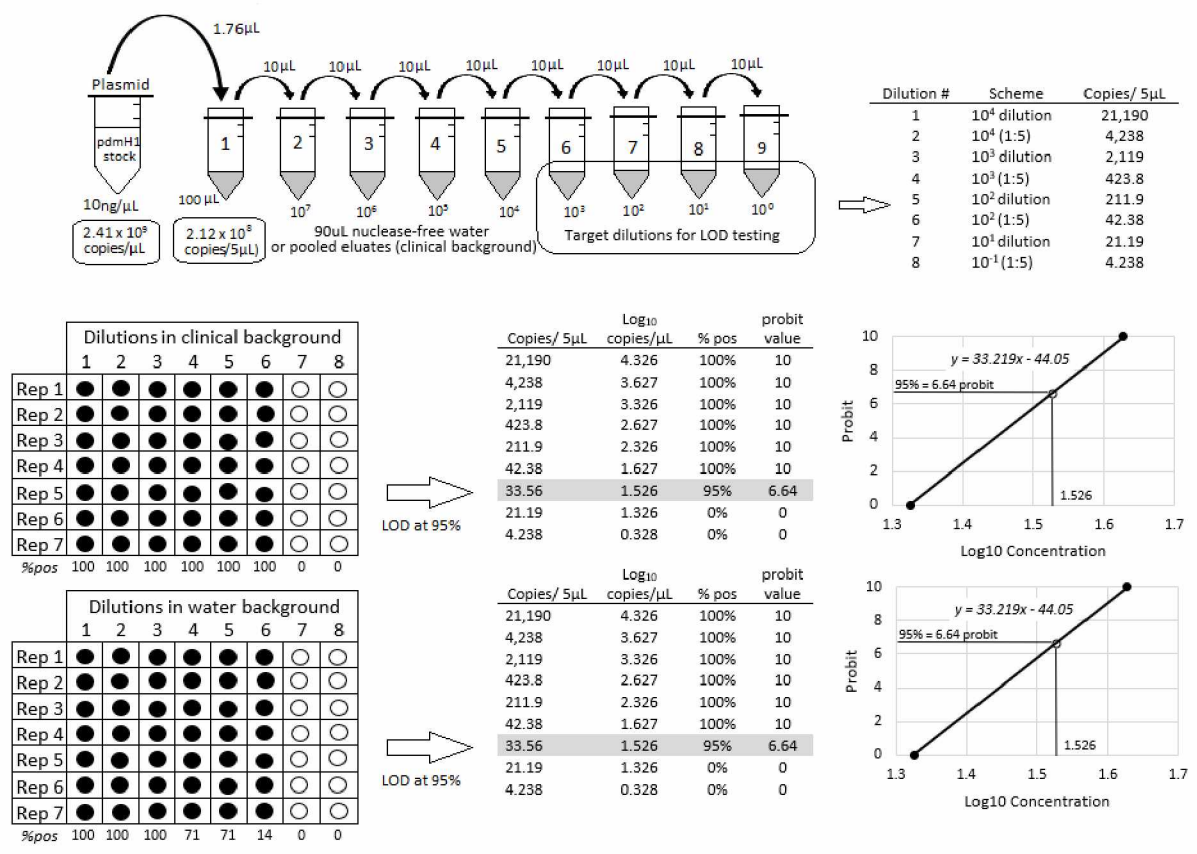


Figure 1.B-4: Influenza A/H1N1 singleplex real-time PCR assay limit of detection determination using probit values

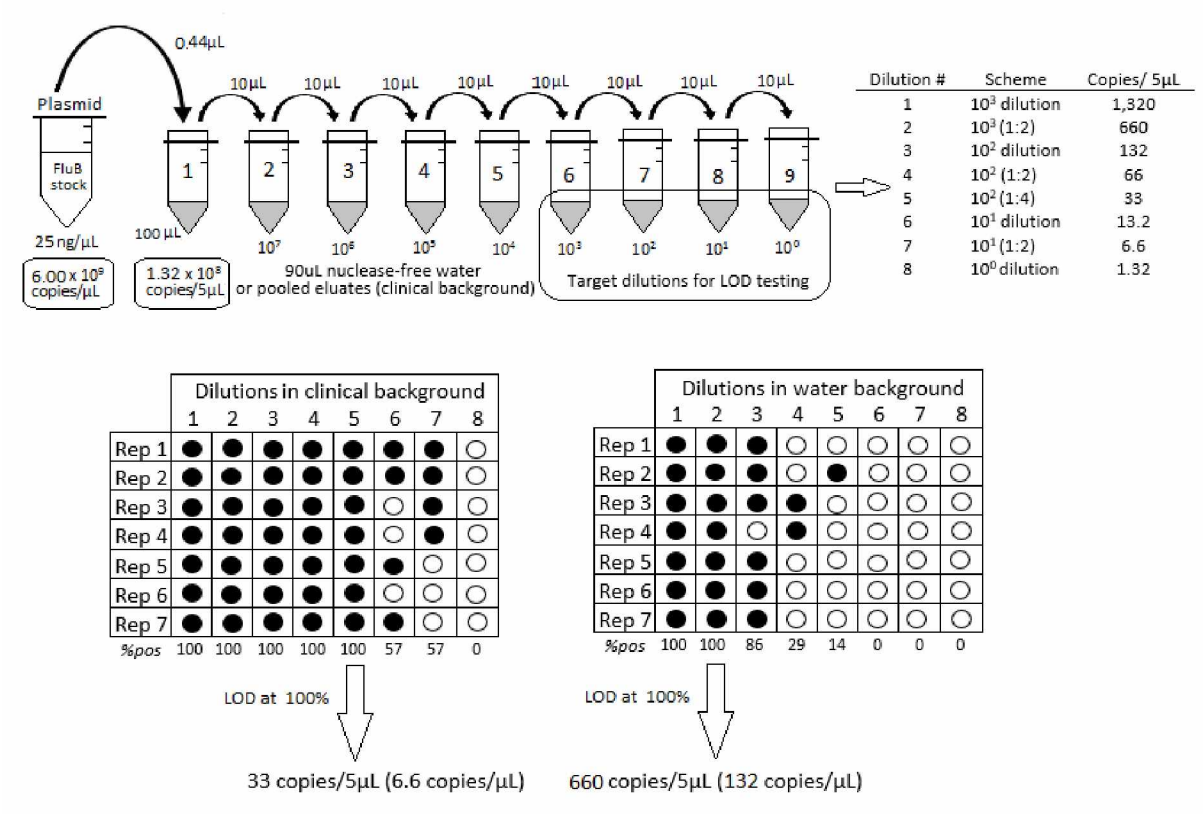


Figure 1.B-5: Influenza B singleplex real-time PCR assay limit of detection determination

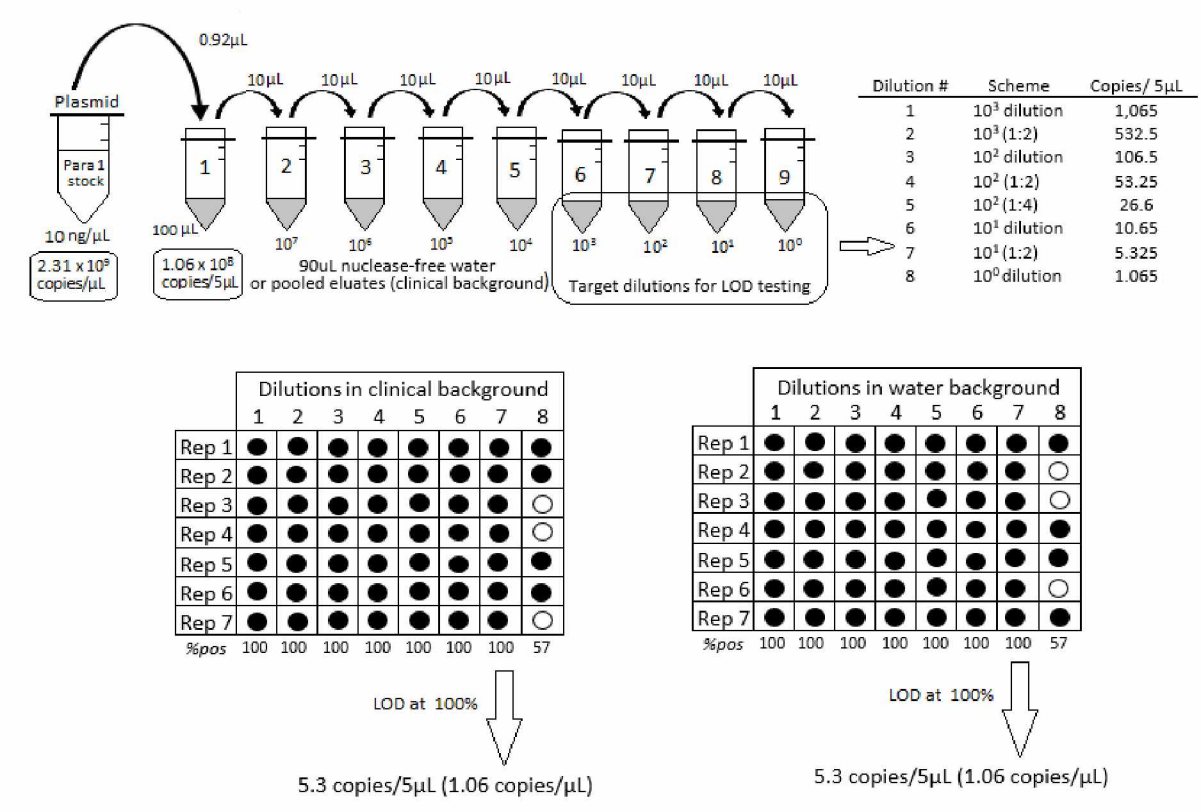


Figure 1.B-6: Parainfluenza 1 singleplex real-time PCR assay limit of detection determination

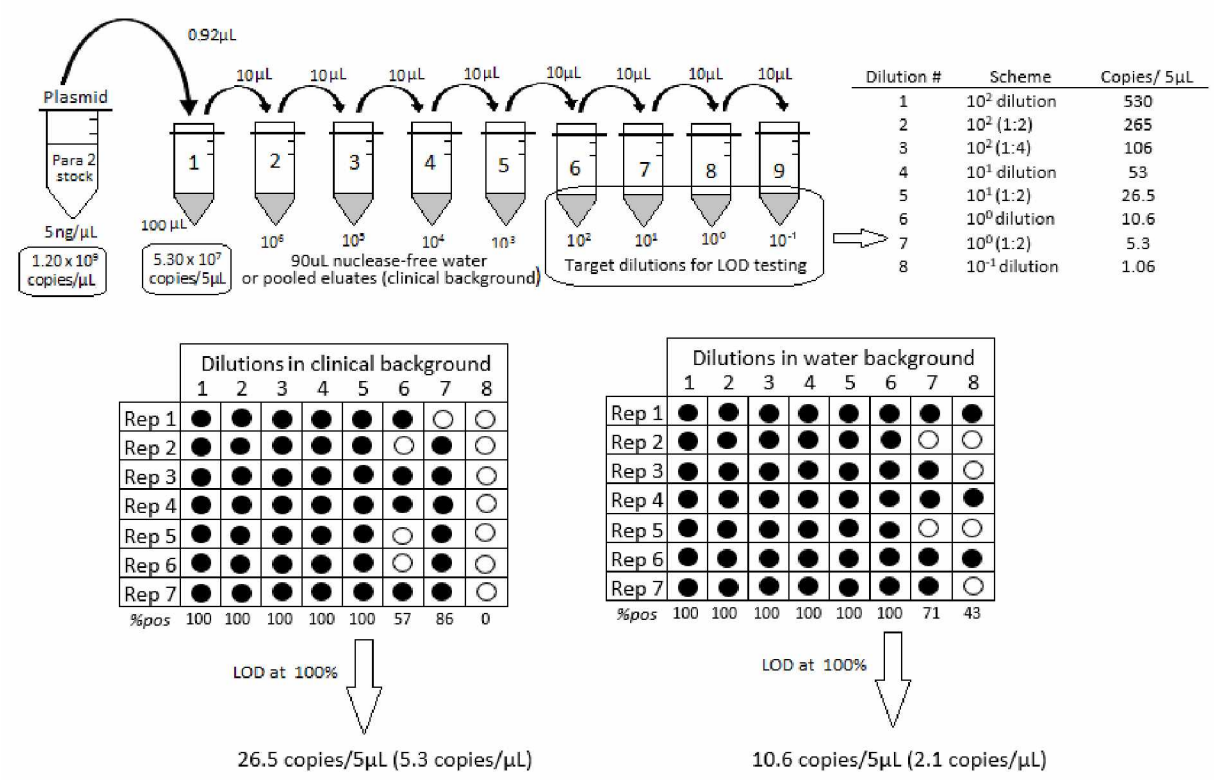


Figure 1.B-7: Parainfluenza 2 singleplex real-time PCR assay limit of detection determination

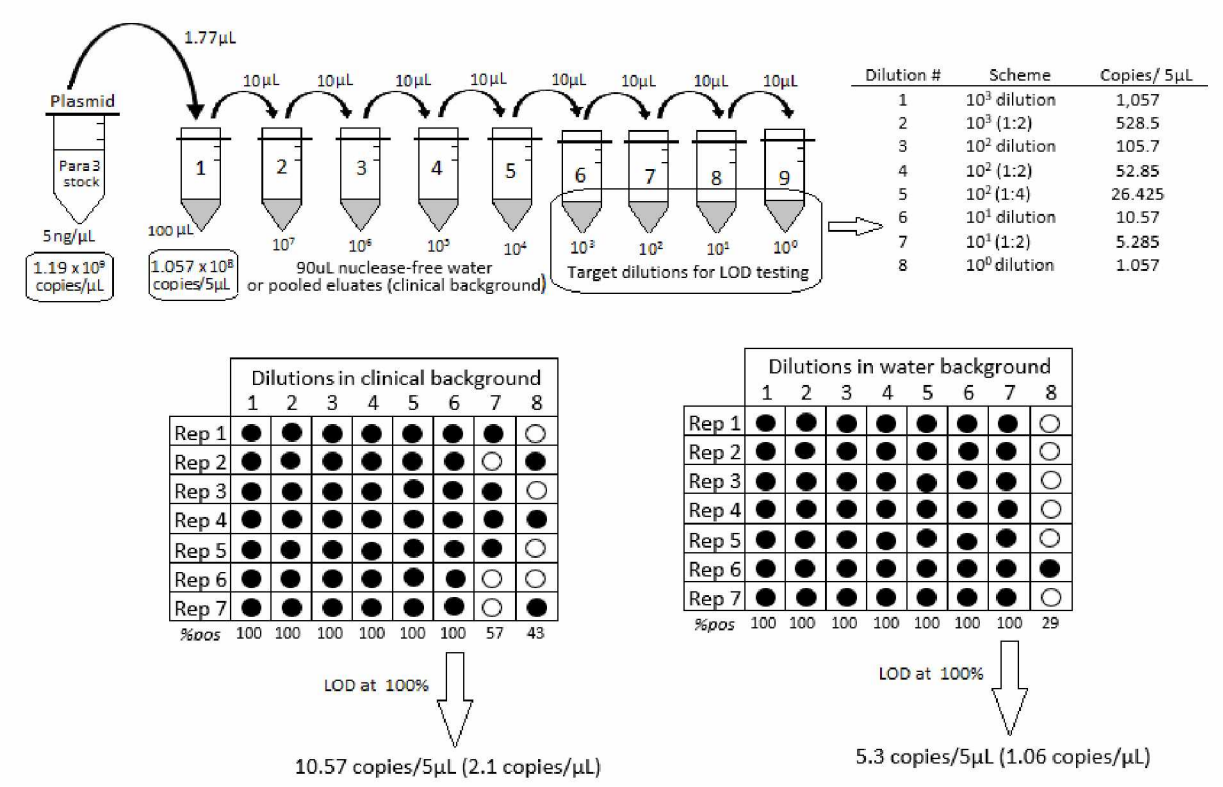


Figure 1.B-8: Parainfluenza 3 singleplex real-time PCR assay limit of detection determination

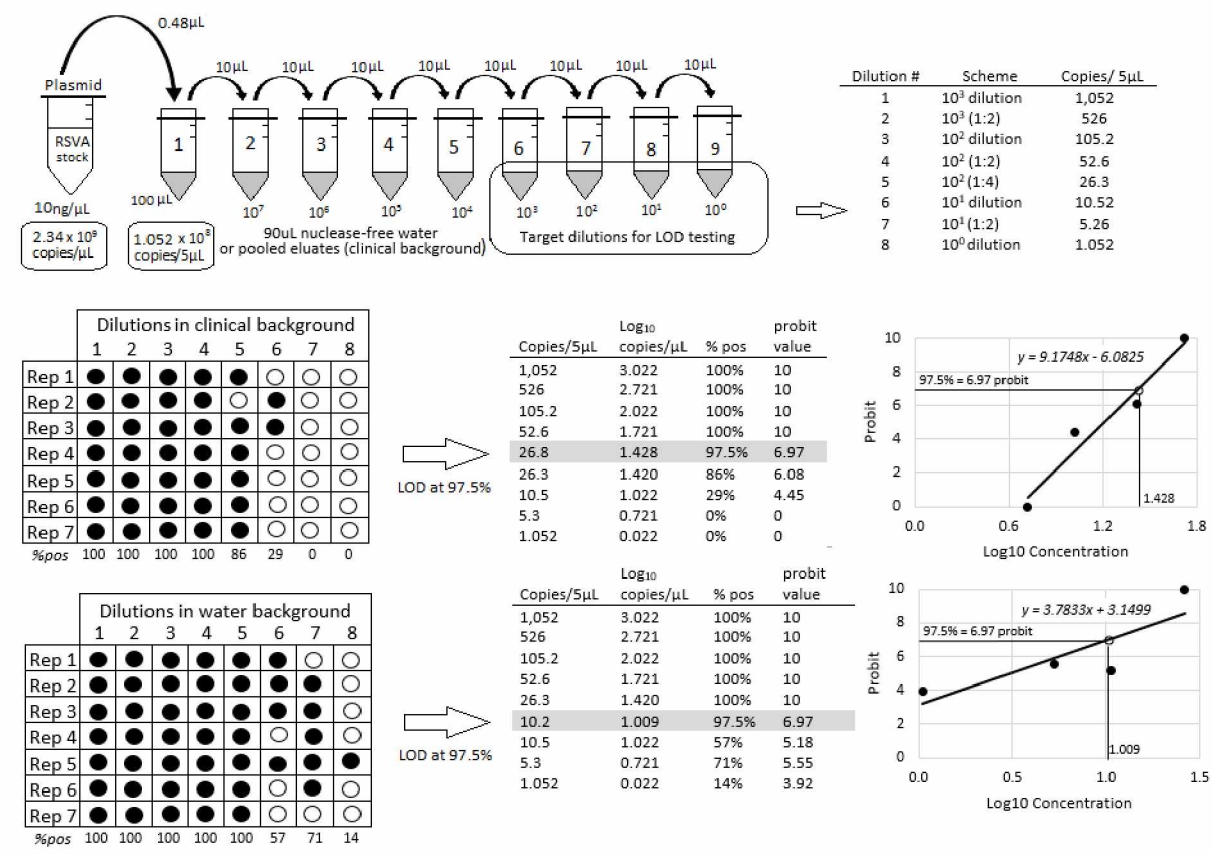


Figure 1.B-9: RSVA singleplex real-time PCR assay limit of detection determination using probit values

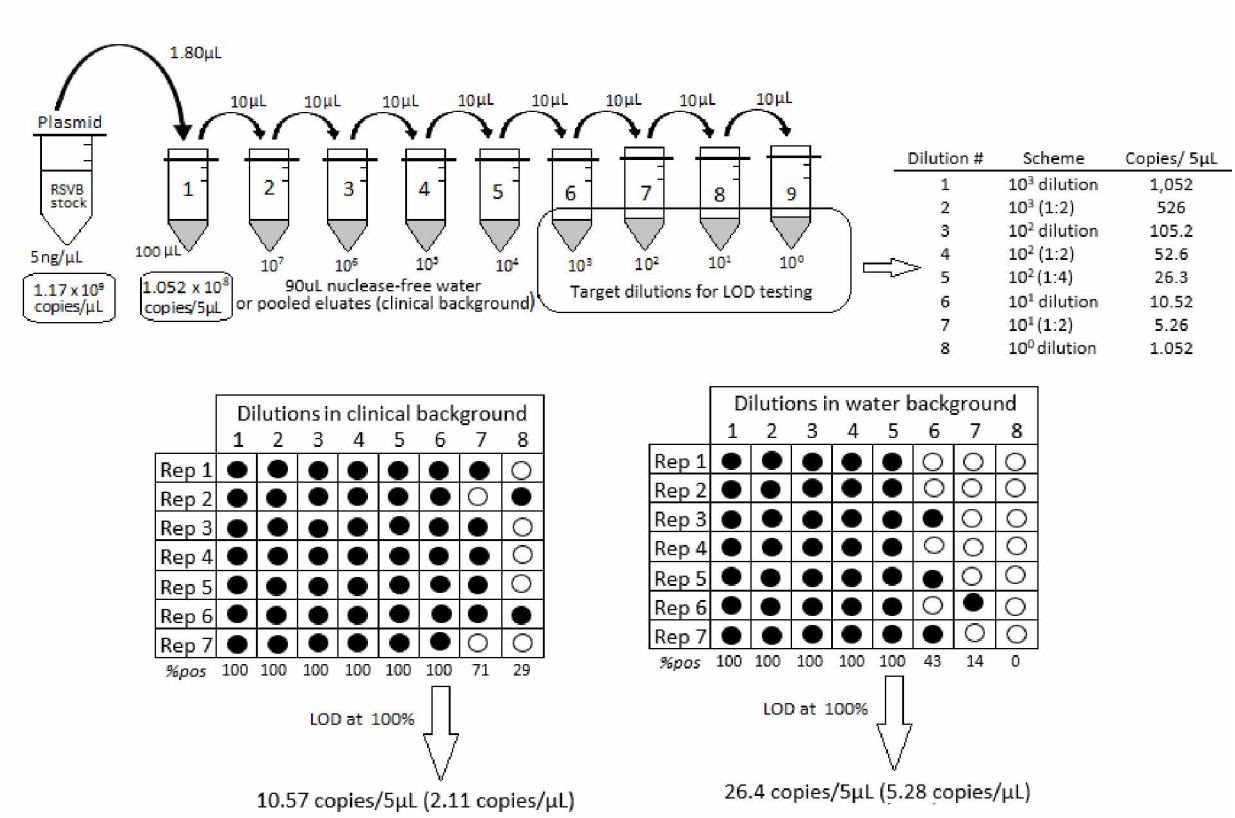


Figure 1.B-10: RSVB singleplex real-time PCR assay limit of detection determination

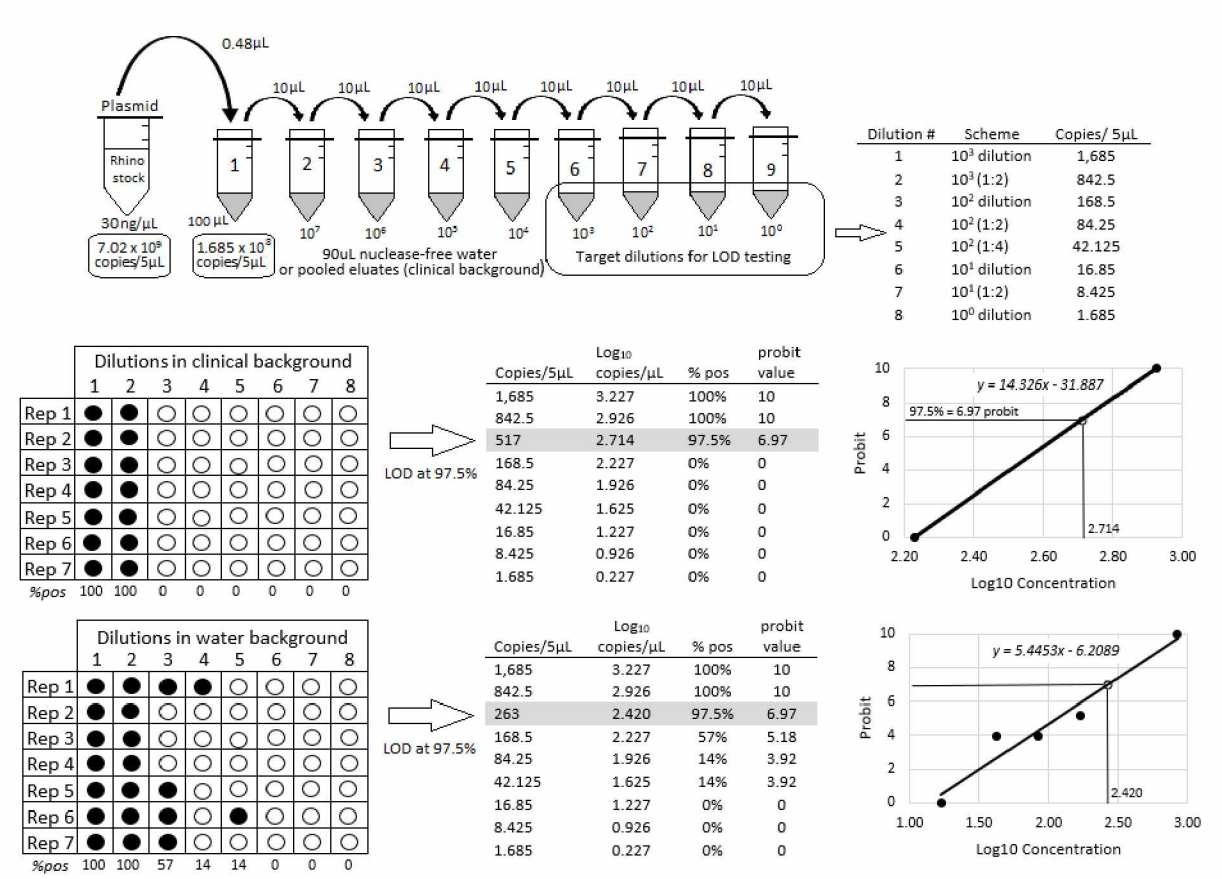
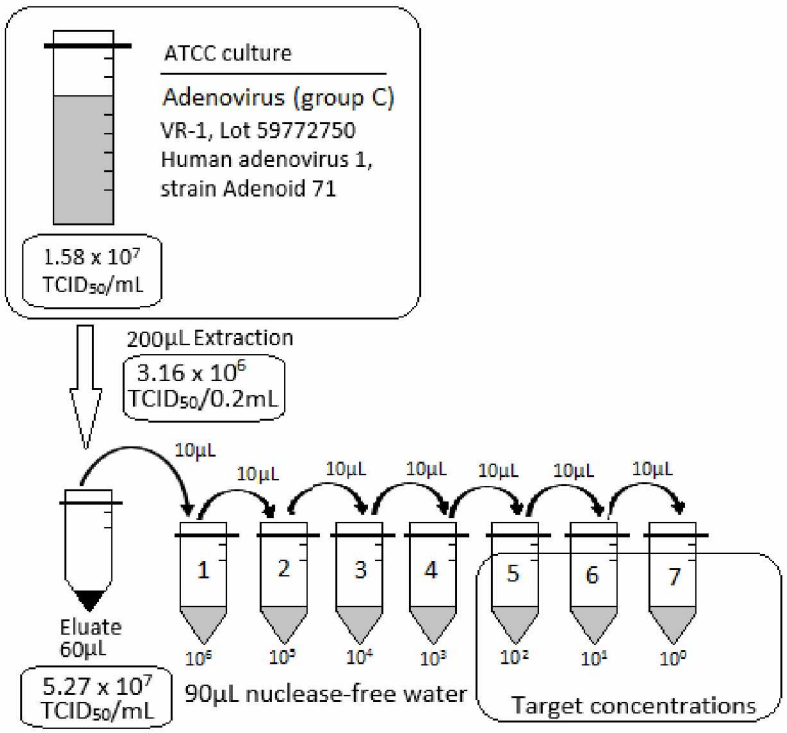


Figure 1.B-11: Rhinovirus singleplex real-time PCR assay limit of detection determination using probit values

Appendix 1.C

Converting TCID dilutions to copy number equivalents

The purpose of this appendix is to show the laboratory methods used to convert tissue culture infectious dose (TCID) concentrations to copy numbers. The same ATCC cultures used to establish the LOD for the GenMark multiplex assays in the FDA trial were used to perform these conversion experiments. ATCC cultures were extracted and various dilutions encompassing or, in the case of influenza B, approaching the reported LOD were targeted via qPCR using plasmids described in Appendix 1.A. The figures in this appendix help show how the concentrations were tracked throughout the process. Averages and standard deviations were calculated from three replicates at each $TCID_{50}$/mL target dilution. These were used to produce a low, average, and high trend line from which to estimate copy numbers for reported TCID LODs. Averages and standard deviations were calculated to produce a final copy number value (per 5μL) at the reported GenMark LOD. Once converted, the copy number value was directly compared to the limits of detection (LOD) established for the singleplex assays in clinical background (Appendix 1.B). Copy number differences between the singleplex and multiplex (GenMark) assays are summarized in the table at the bottom of each figure.



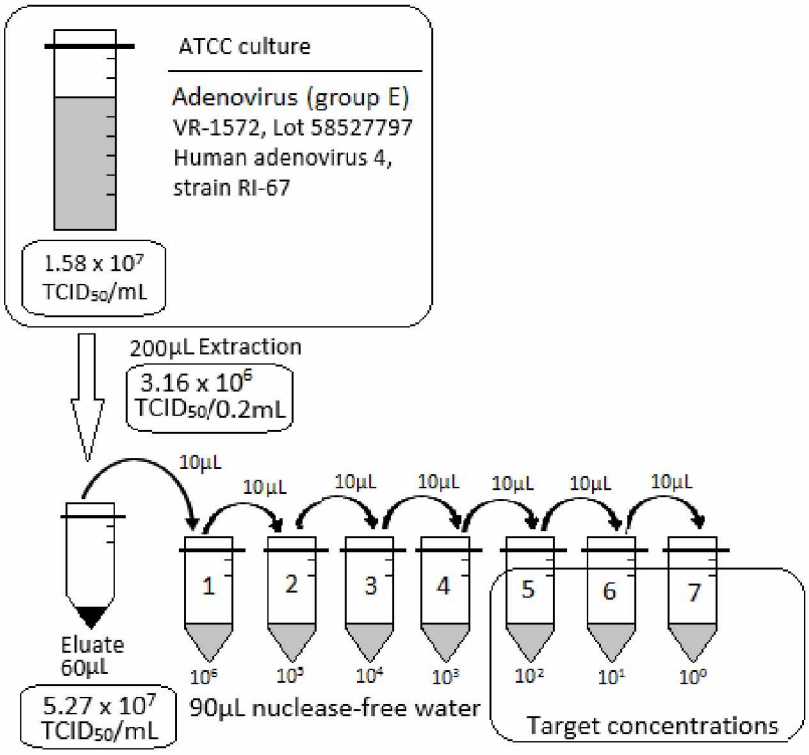
TCID conversion to copy numbers :

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg + SD)
5.27 x 10^2 $TCID_{50}$/mL	5,220 ± 660	4,560	5,220	5,880
5.27 x 10^1 $TCID_{50}$/mL	906 ± 822	85	906	1,729
5.27 x 10^0 $TCID_{50}$/mL	46 ± 13	33	46	59
	Equation of the line	y = 8.9844x - 192.57 R^2 = 0.995	y = 9.5838x + 188.71 R^2 = 0.995	y = 10.183x + 569.98 R^2 = 0.961
	If x = 8.89 x 10^1	606	1,041	1,476
	Avg ± SD (copies/5µL)	1,041 ± 435		

Summary - Limit of detection differences between adenovirus C assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	606	7.92	598	2.78
Average	1,041	7.92	1,032	3.01
High	1,476	7.92	1,468	3.17
Avg ± SD (copies/5µL)	1,041 ± 435	7.92 ± 0	1,033 ± 435	2.99 ± 0.2

Figure 1.C-1: Adenovirus (group C, VR-1) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 8.89 x 10^1 $TCID_{50}$/mL.



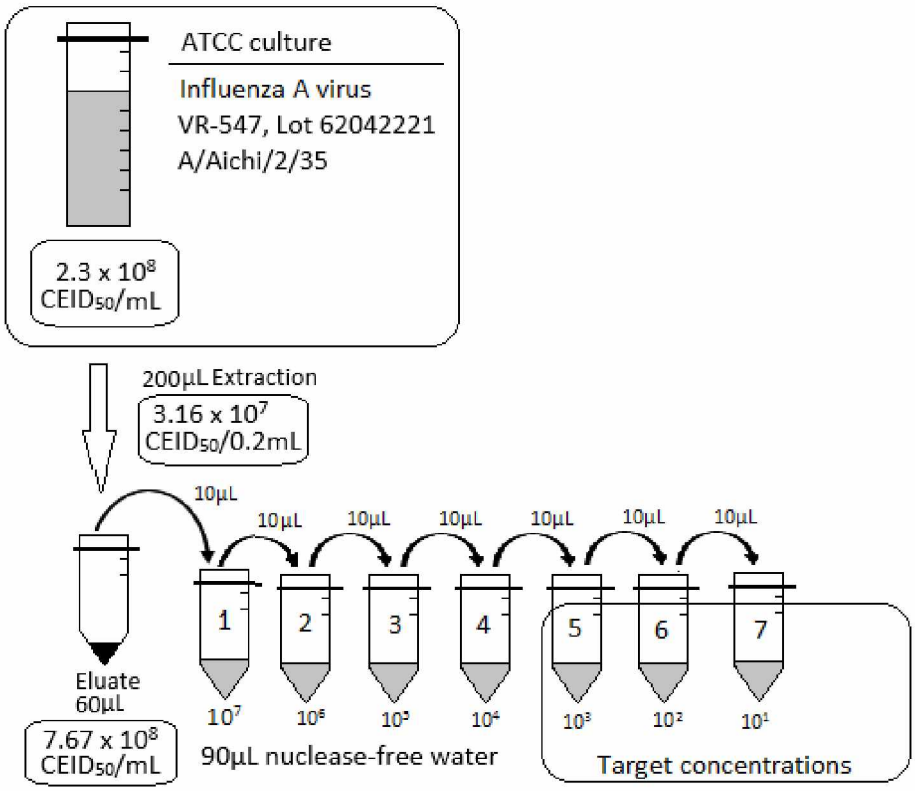
TCID conversion to copy numbers:

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg + SD)
5.27 x 10^2 $TCID_{50}$/mL	78,711 ± 5,290	73,421	78,711	84,001
5.27 x 10^1 $TCID_{50}$/mL	9,059 ± 945	8,114	9,059	10,004
5.27 x 10^0 $TCID_{50}$/mL	807 ± 86	721	807	893
	Equation of the line	y = 157.96x + 831.39 R^2 = 0.999	y = 148.32x + 605 R^2 = 0.999	y = 138.67x + 378.61 R^2 = 0.999
	If x = 8.89 x 10^1	3,327	2,948	2,570
	Avg ± SD (copies/5µL)	2,948 ± 379		

Summary - Limit of detection differences between adenovirus B/E assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	2,570	7.92	2,562	3.41
Average	2,949	7.92	2,941	3.47
High	3,327	7.92	3,319	3.52
Avg ± SD (copies/5µL)	2,948 ± 379	7.92 ± 0	2,941 ± 379	3.47 ± 0.6

Figure 1.C-2: Adenovirus (group E, VR-1572) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 1.58 x 10^1 $TCID_{50}$/mL.



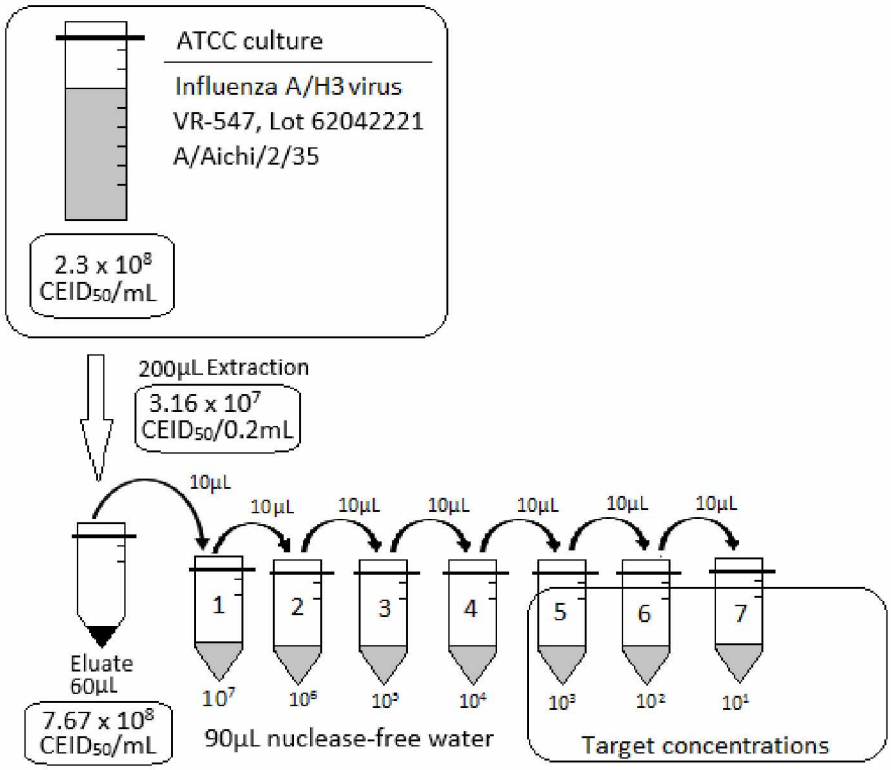
TCID conversion to copy numbers:

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
7.67 x 10^3 $CEID_{50}$/mL	33,686 ± 4,308	29,378	33,686	37,994
7.67 x 10^2 $CEID_{50}$/mL	2,823 ± 335	2,488	2,823	3,158
7.67 x 10^1 $CEID_{50}$/mL	403 ± 26	377	403	429
	Equation of the line	y = 3.8501x - 178.67 R^2=0.999	y = 4.4188x - 236.06 R^2=0.999	y = 4.9874x - 293.44 R^2=0.999
	If x = 1.58 x 10^3	5,916	6,759	7,602
	Avg ± SD (copies/5µL)	6,759 ± 843		

Summary – Limit of detection differences between influenza A (generic) assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	5,916	26.5	5,890	3.77
Average	6,759	26.5	6,732	3.83
High	7,602	26.5	7,575	3.88
Avg ± SD (copies/5µL)	6,759 ± 843	26.5 ± 0	6,732 ± 843	3.83 ± 0.05

Figure 1.C-3: Influenza A (VR-547) TCID concentration conversion to copy numbers, CEID = chicken egg infectious dose, GenMark RVP LOD to convert to copy numbers = 1.58 x 10^3 $CEID_{50}$/mL.



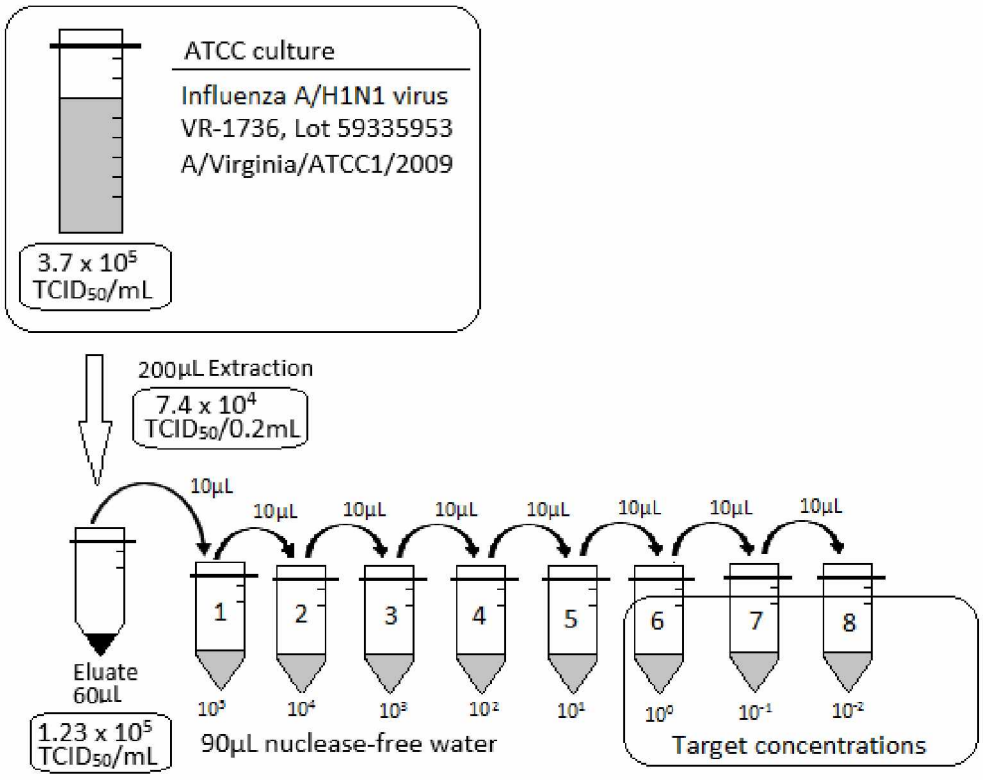
TCID conversion to copy numbers:

Tested concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
7.67 x 10^3 $CEID_{50}$/mL	2.7 ± 1.2	1.5	2.7	3.9
7.67 x 10^2 $CEID_{50}$/mL	3.7 ± 1.9	1.8	3.7	5.6
7.67 x 10^1 $CEID_{50}$/mL	3.3 ± 1.2	2.1	3.3	4.5
	Equation of the line	y = -0.00006x + 1.98 R^2 = 0.818	y = -0.0004x + 4.86 R^2 = 0.818	y = -0.0001x + 5.08 R^2 = 0.818
	If x = 1.58 x 10^3	1.9	4.2	4.9
	Avg ± SD (copies/5µL)	3.7 ± 1.6		

Summary – Limit of detection differences between influenza A/H3 assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	2.1	80	77.9	1.89
Average	3.7	80	76.3	1.88
High	5.3	80	74.7	1.87
Avg ± SD (copies/5µL)	3.7 ± 1.6	80 ± 0	76.3 ± 1.6	1.88 ± 0.01

Figure 1.C-4: Influenza A/H3 (VR-547) TCID concentration conversion to copy numbers, CEID = chicken egg infectious dose. The LOD for the multiplex GenMark influenza A/H3 assay exceeded the LOD of the qPCR assay and conversions are considered estimates based on continued trend lines beyond the LOD of the qPCR assay. GenMark RVP LOD to convert to copy numbers = 1.58 x 10^3 $CEID_{50}$/mL.



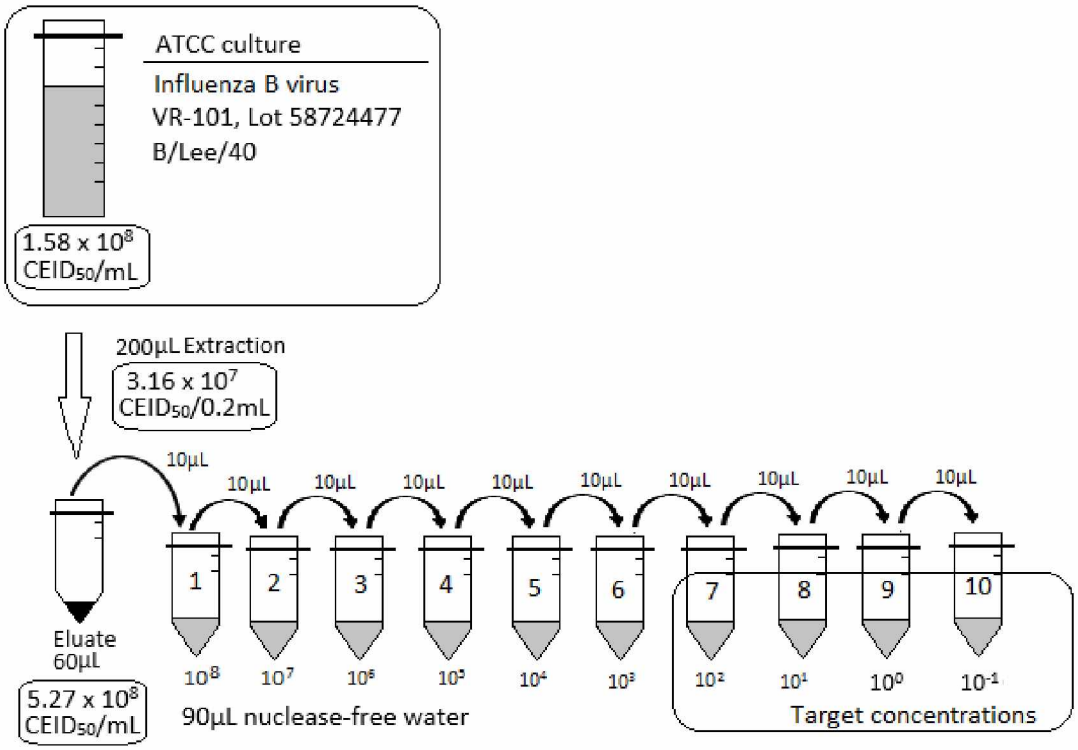
TCID conversion to copy numbers:

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
1.23 x 10^0 $TCID_{50}$/mL	9,705 ± 2,241	7,464	9,705	11,946
1.23 x 10^{-1} $TCID_{50}$/mL	512 ± 131	381	512	643
1.23 x 10^{-2} $TCID_{50}$/mL	57 ± 24	33	57	81
	Equation of the line	y = 6,222.4x – 205.83 R^2= 0.998	y = 8,077.7x – 251.5 R^2= 0.998	y = 9,933x – 297.17 R^2=0.998
	If x = 1.05 x 10^{-1}	448	597	746
	Avg ± SD (copies/5µL)	597 ± 149		

Summary – Limit of detection differences between influenza A/H1 assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	448	16.9	431	2.63
Average	597	16.9	580	2.76
High	746	16.9	729	2.86
Avg ± SD (copies/5µL)	597 ± 149	17 ± 0	580 ± 149	2.75 ± 0.11

Figure 1.C-5: Influenza A/H1N1 (VR-1736) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 1.05 x 10^{-1} $TCID_{50}$/mL.



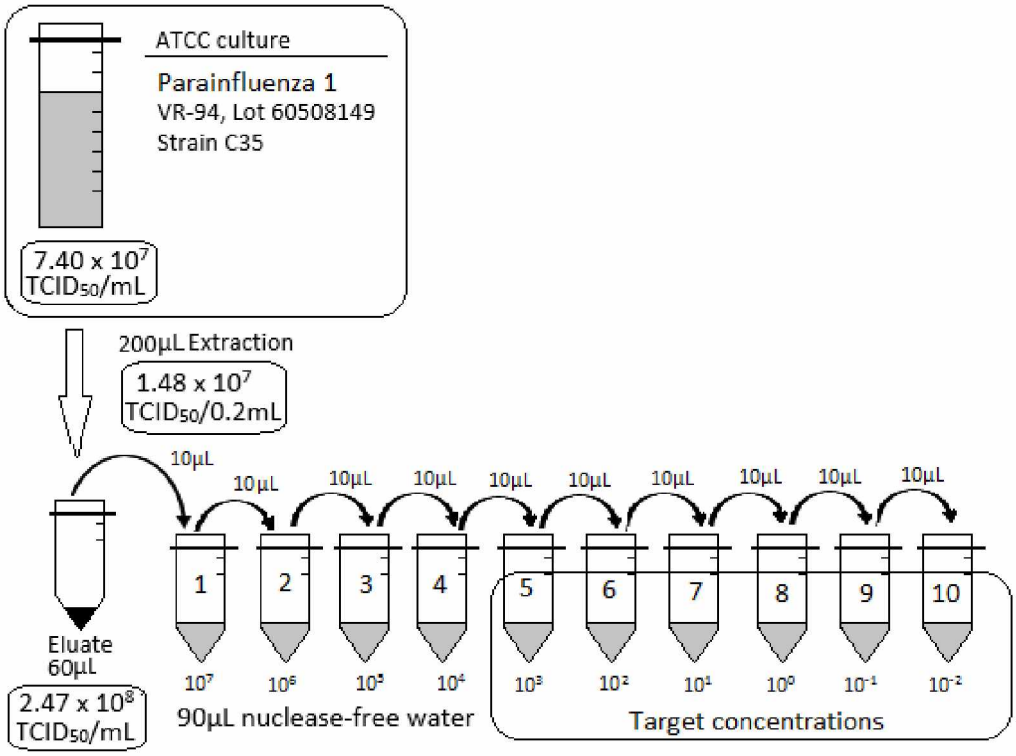
TCID conversion to copy numbers:

Tested concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
5.27 x 10^2 $CEID_{50}$/mL	1,131 ± 121	1,010	1,131	1,252
5.27 x 10^1 $CEID_{50}$/mL	130 ± 10	120	130	140
5.27 x 10^0 $CEID_{50}$/mL	8 ± 10	-2	8	18
5.27 x 10^{-1} $CEID_{50}$/mL	1 ± 1.7	-0.7	1	2.7
	Equation of the line	y = 1.9174x + 1.1646 R^2=0.999	y = 2.1405x + 4.1889 R^2=0.999	y = 2.3635x + 7.2132 R^2=0.999
	If x = 3.16 x 10^{-1}	3.4	7.3	11.2
	Avg ± SD (copies/5µL)	4.9 ± 3.1		

Summary - Limit of detection differences between influenza B assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	1.8	33	31.2	1.49
Average	4.9	33	28.1	1.45
High	8.0	33	25.0	1.40
Avg ± SD (copies/5µL)	4.9 ± 3.1	33 ± 0	28.1 ± 2.5	1.45 ± 0.04

Figure 1.C-6: Influenza B (VR-101) TCID concentration conversion to copy numbers, CEID = chicken egg infectious dose. The LOD for the multiplex GenMark influenza B assay exceeded the LOD of the qPCR assay and conversions are considered estimates based on continued trend lines beyond the LOD of the qPCR assay. GenMark RVP LOD to convert to copy numbers = 3.16 x 10^{-1} $CEID_{50}$/mL.



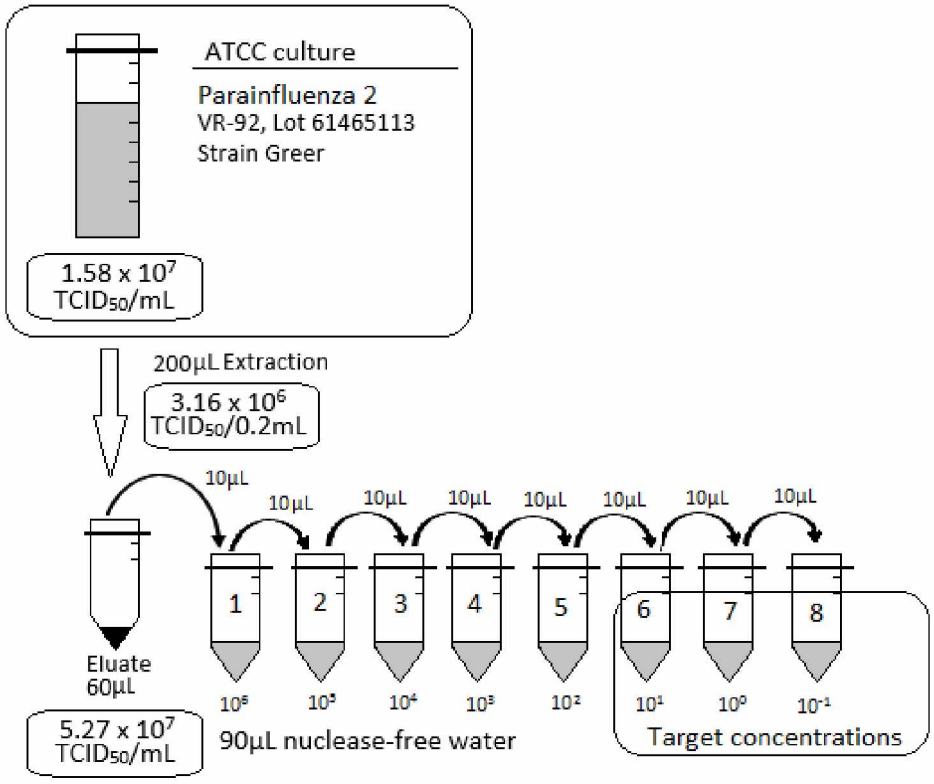
TCID conversion to copy numbers:

Tested concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
2.47 x 10^2 $TCID_{50}$/mL	1,747 ± 82	1,665	1,747	1,829
2.47 x 10^1 $TCID_{50}$/mL	205 ± 5	200	205	210
2.47 x 10^0 $TCID_{50}$/mL	29 ± 13	16	29	42
2.47 x 10^{-1} $TCID_{50}$/mL	0	0	0	0
2.47 x 10^{-2} $TCID_{50}$/mL	4 ± 4	0	4	8
	Equation of the line	y = 6.7221x + 7.2377	y = 7.0378x + 10.707	y = 7.3535x + 14.177
		R^2 = 0.999	R^2 = 0.999	R^2 = 0.999
	If x = 2.81 x 10^{-2}	7.4	10.9	14.4
	Avg ± SD (copies/5µL)	10.9 ± 3.5		

Summary - Limit of detection differences between parainfluenza 1 assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	7.4	5.3	2.1	0.33
Average	10.9	5.3	5.6	0.75
High	14.4	5.3	9.1	0.96
Avg ± SD (copies/5µL)	10.9 ± 3.5	5.3 ± 0	5.6 ± 3.5	0.68 ± 0.32

Figure 1.C-7: Parainfluenza 1 (VR-94) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 2.81 x 10^{-2} $TCID_{50}$/mL.



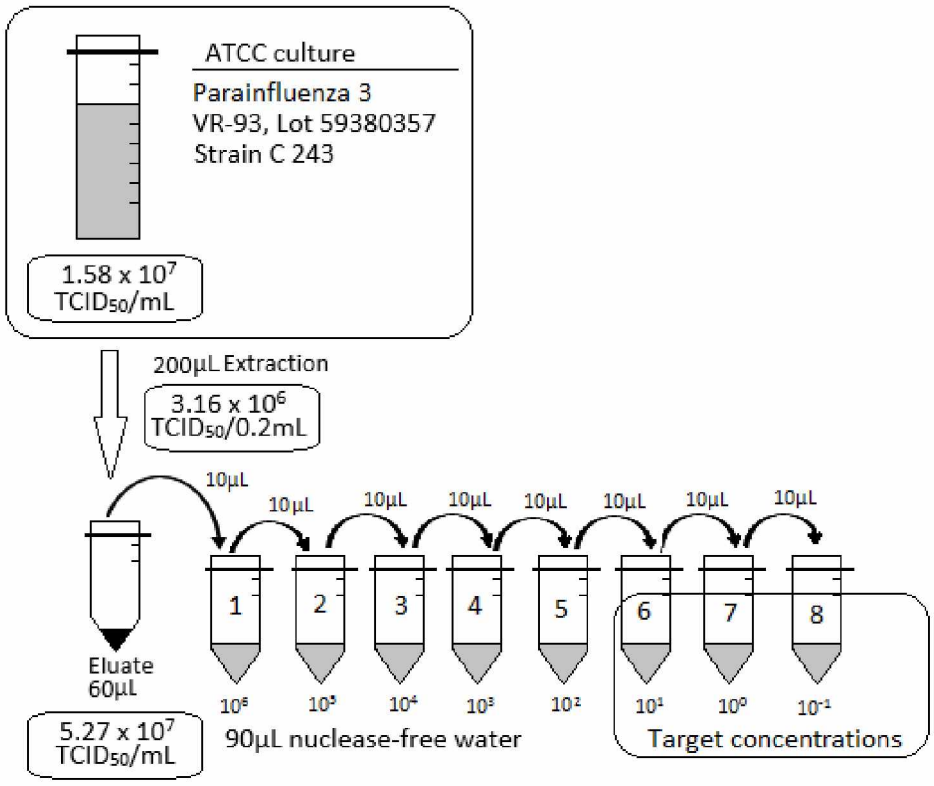
TCID conversion to copy numbers:

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
5.27 x 10^1 $TCID_{50}$/mL	817 ± 116	701	817	933
5.27 x 10^0 $TCID_{50}$/mL	51 ± 5	46	51	56
5.27 x 10^{-1} $TCID_{50}$/mL	7 ± 3	4	7	10
	Equation of the line	y = 13.542x - 13.722 R^2=0.999	y = 15.779x – 16 R^2=0.998	y = 18.015x – 18.278 R^2=0.998
	If x = 2.81 x 10^0	24.3	28.3	32.3
	Avg ± SD (copies/5µL)	28.3 ± 4		

Summary - Limit of detection differences between parainfluenza 2 assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	24.3	26.5	2.2	0.34
Average	28.3	26.5	1.8	0.26
High	32.3	26.5	5.8	0.77
Avg ± SD (copies/5µL)	28.3 ± 4	26.5 ± 0	3.3 ± 2.2	0.46 ± 0.27

Figure 1.C-8: Parainfluenza 2 (VR-92) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 2.81 x 10^0 $TCID_{50}$/mL.



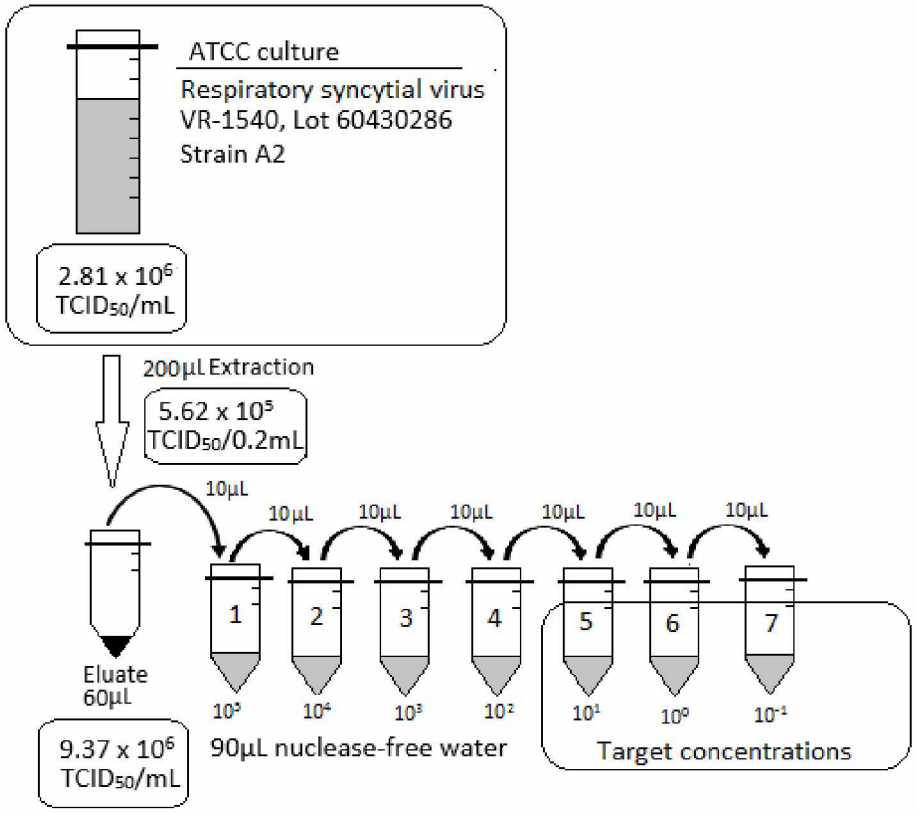
TCID conversion to copy numbers:

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
5.27 x 10^1 $TCID_{50}$/mL	1,375 ± 176	1,199	1,375	1,551
5.27 x 10^0 $TCID_{50}$/mL	132 ± 36	96	132	168
5.27 x 10^{-1} $TCID_{50}$/mL	32 ± 12	20	32	44
	Equation of the line	y = 22.864x - 7.5	y = 25.93x + 7.3889	y = 28.996x + 22.278
		R^2 = 0.999	R^2 = 0.999	R^2 = 0.999
	If x = 2.81 x 10^1	635	736	837
	Avg ± SD (copies/5µL)	736 ± 101		

Summary - Limit of detection differences between parainfluenza 3 assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	635	10.6	624	2.80
Average	736	10.6	725	2.86
High	837	10.6	826	2.92
Avg ± SD (copies/5µL)	736 ± 101	10.6 ± 0	725 ± 101	2.86 ± 0.06

Figure 1.C-9: Parainfluenza 3 (VR-93) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 2.81 x 10^1 $TCID_{50}$/mL.



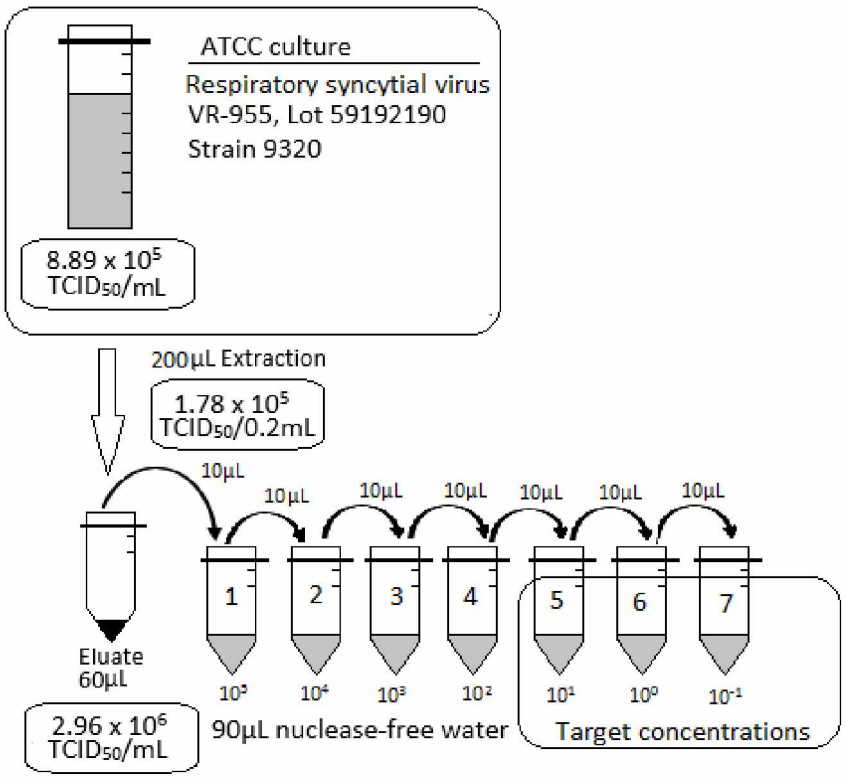
TCID conversion to copy numbers:

Target concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
9.37 x 10^1 $TCID_{50}$/mL	38,523 ± 1,544	36,979	38,523	40,067
9.37 x 10^0 $TCID_{50}$/mL	5,304 ± 362	4,942	5,304	5,666
9.37 x 10^{-1} $TCID_{50}$/mL	553 ± 38	515	553	591
	Equation of the line	y = 409.59x + 596.72 R^2 = 0.999	y = 415.47x + 950.41 R^2 = 0.998	y = 383.85x + 910.65 R^2 = 0.998
	If x = 2.81 x 10^0	1,748	2,118	1,989
	Avg ± SD (copies/5µL)	1,952 ± 188		

Summary - Limit of detection differences between RSVA assays:

Copy number variance	GenMark LOD	Singlelex LOD	Difference	Log
Low	1,764	23.4	1,741	3.24
Average	1,952	23.4	1,929	3.29
High	2,140	23.4	2,117	3.33
Avg ± SD (copies/5µL)	1,952 ± 188	23.4 ± 0	1,929 ± 154	3.28 ± 0.03

Figure 1.C-10: RSVA (VR-1540) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 2.81 x 10^0 $TCID_{50}$/mL.



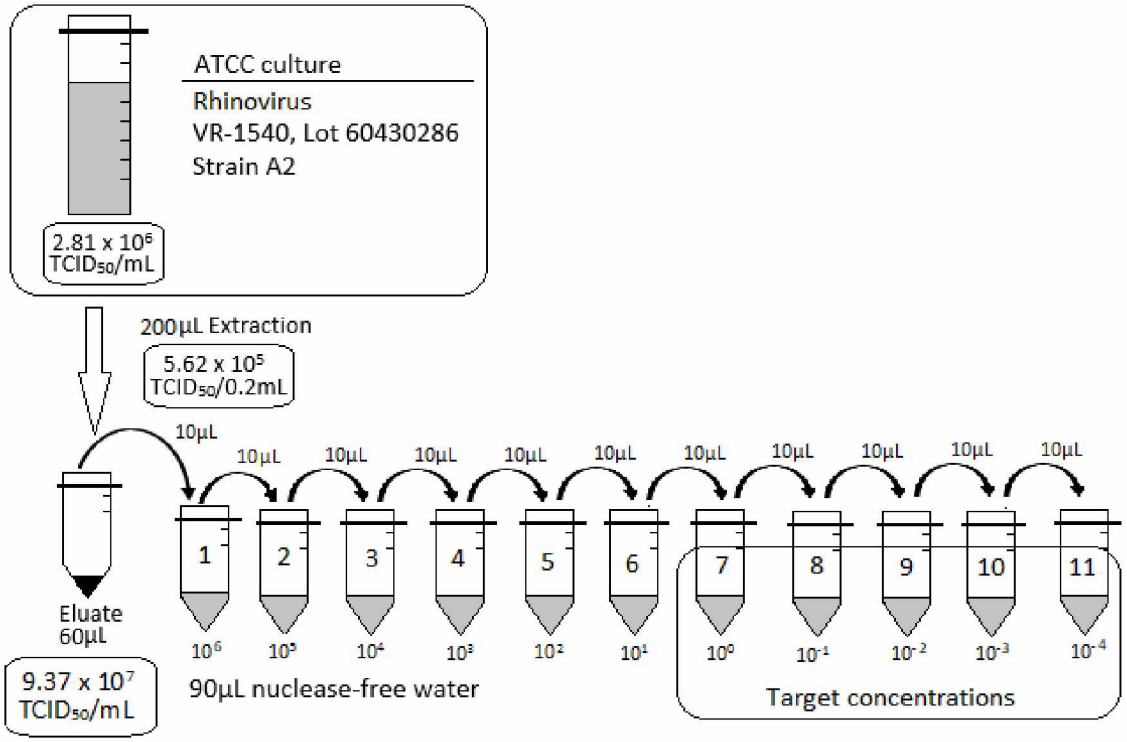
TCID conversion to copy numbers:

Tested concentrations	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
2.96 x 10^1 $TCID_{50}$/mL	8,376 ± 1,317	7,059	8,376	9,693
2.96 x 10^0 $TCID_{50}$/mL	1,164 ± 88	1,076	1,164	1,252
2.96 x 10^{-1} $TCID_{50}$/mL	61 ± 9	52	61	70
	Equation of the line	y = 233.22x + 174.72 R^2 = 0.997	y = 278.47x + 150.56 R^2 = 0.999	y = 323.71x + 126.39 R^2 = 0.999
	If x = 1.58 x 10^0	543	591	638
	Avg ± SD (copies/5µL)	591 ± 47		

Summary – Limit of detection differences between RSVB assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	543	11	532	2.73
Average	591	11	580	2.76
High	638	11	627	2.80
Avg ± SD (copies/5µL)	591 ± 47	11 ± 0	580 ± 47	2.76 ± 0.04

Figure 1.C-11: RSVB (VR-955) TCID concentration conversion to copy numbers, GenMark RVP LOD to convert to copy numbers = 1.58 x 10^0 $TCID_{50}$/mL.



TCID conversion to copy numbers:

Target concentration	Copies/5µL Avg ± SD	Low (Avg - SD)	Avg	High (Avg +SD)
9.37 x 10^{0} $TCID_{50}$/mL	4,829 ± 1,170	3,659	4,829	5,999
9.37 x 10^{-1} $TCID_{50}$/mL	236 ± 228	7.9	236.0	464.1
9.37 x 10^{-2} $TCID_{50}$/mL	2.17 ± 1.92	0.3	2.2	4.1
9.37 x 10^{-3} $TCID_{50}$/mL	5.87 ± 8.13	-2.3	5.9	14.0
9.37 x 10^{-4} $TCID_{50}$/mL	0.09 ± 0.14	-0.1	0.1	0.2
	Equation of the line	y = 397.9x - 95.54 R^2 = 0.991	y = 520.93x - 70.051 R^2 = 0.998	y = 643.98x - 44.615 R^2 = 0.999
	If x = 1.58 x 10^{-3}	-94.9	70.9	45.6
	Avg ± SD (copies/5µL)	7 ± 89		

Summary - Limit of detection differences between rhinovirus assays:

Copy number variance	GenMark LOD	Singleplex LOD	Difference	Log
Low	-82	517	599	2.78
Average	7	517	510	2.71
High	97	517	420	2.62
Avg ± SD (copies/5µL)	7 ± 89	517 ± 0	510 ± 89	2.70 ± 0.08

Figure 1.C-12: Rhinovirus (VR-1540) TCID concentration conversion to copy numbers. The LOD for the multiplex GenMark rhinovirus assay exceeded the LOD of the qPCR assay and conversions are considered estimates based on continued trend lines beyond the LOD of the qPCR assay. GenMark RVP LOD to convert to copy numbers = 1.58 x 10^{-3} $TCID_{50}$/mL.

Chapter 2 BioRad BioPlex® HIV Ag-Ab assay: Incidence of false positivity in a low-prevalence population and its effects on the current HIV testing algorithm[2]

2.1 Abstract

The BioPlex® HIV Ag-Ab assay, unlike other HIV 1/2 antigen/antibody immunoassays, is capable of differentiating positive HIV-1 antibodies (Groups M and O) from HIV-2 antibodies and/or HIV-1 p24 antigen in a single test. The Alaska State Virology Laboratory (ASVL) adopted the BioPlex® HIV Ag-Ab assay early 2017 and can report on its performance in terms of false positivity in a low-prevalence population and its effects on the current HIV testing algorithm recommended by the Centers for Disease Control and Prevention (CDC). Specimens received between March 2017 and August 2018 were screened using the BioPlex® HIV Ag-Ab assay. Specimens screening positive for HIV antibodies or antigen were further confirmed using the Geenius™ HIV 1/2 Supplemental Assay and/or HIV RNA testing. Of the 12,338 sera screened for HIV, 35 specimens were positive on the BioPlex® HIV Ag-Ab assay. Only 22 of the specimens were able to be confirmed by the Geenius™ HIV 1/2 Supplemental Assay and were considered to be truly positive (PPV, 62.9%). RNA was not detected in these cases suggesting initial false positivity on the BioPlex® HIV Ag-Ab assay. True positive results had index values (IDX) of >180 whereas false positive IDX's were between 1 and 4, with the exception of one specimen. We suggest that specimens demonstrating positivity with low IDX

[2] Parker J, Carrasco A, Chen J. *BioRad BioPlex® HIV Ag-Ab assay: Incidence of false positivity in a low-prevalence population and its effects on the current HIV testing algorithm*. Journal of Clinical Virology, 2019. **116**: p. 1-3.

values ≤4 on the BioPlex® HIV Ag-Ab assay proceed directly to RNA testing, essentially bypassing supplemental antibody confirmation tests, to reduce turnaround time and cost of HIV confirmation.

2.2 Introduction

Broadening the detection window while improving HIV assay sensitivity has made a positive impact on HIV patient management and treatment. However, improved sensitivity in combination with other biological factors and technical issues can lead to false positive HIV results [1-4]. The HIV 1/2 antigen/antibody immunoassays have been shown to produce false positive results, especially in populations with low HIV prevalence [5, 6]. False positives are identified as specimens with low reactivity by initial HIV1/2 antigen/antibody immunoassays, which are unable to be confirmed by subsequent testing using varied methodologies, such as the Bio-Rad Geenius™ HIV-1/2 Supplemental Assay and PCR [4, 7-10].

BioRad's BioPlex® HIV Ag-Ab combination assay is a multiplex flow immunoassay that can simultaneously detect and differentiate HIV-1 p24 antigen, HIV-1 (groups M and O) antibodies, and HIV-2 antibodies in human serum or plasma. It reportedly produces the best analytical sensitivity of HIV-1 p24 antigen on the market (limit of detection, 0.33 IU/mL and 5.2 pg/mL) as well as high specificity in low risk population (99.86%) [11]. The assay has been shown to be effective in identifying early cases of HIV when compared to other automated platforms [12]. The BioPlex® HIV Ag-Ab assay is not widely used in public health settings at this time but offers a clear advantage in terms of reducing labor while improving diagnostics by separating antigen-antibody combination results into individual measurements in a fully automated manner. In this study, we look at the false positivity rate and positive predictive value

when using the BioPlex® HIV Ag-Ab combination assay in a low HIV prevalence population and provide insight on its effects on the recommended HIV testing algorithm.

2.3 Methods and Materials

A total of 12,338 sera were collected from patients ranging from age 2 to 92 (median age = 29, interquartile range = 23 to 38 years old) between March 2017 and August 2018 from various regions across Alaska and sent to the ASVL for surveillance purposes.

Specimens were screened using the BioPlex® HIV Ag-Ab assay on the BioPlex® 2200 instrument. Results are expressed as an index value (IDX) describing the measured RFI (relative fluorescent intensity unit) as a ratio to the cut-off value for each particular bead type. Four separate results were generated for each multiplexed test, where IDX values of > 1 were considered reactive: 1) HIV Ag/Ab combo, undifferentiated, 2) HIV-1 antibody, 3) HIV-1 antigen, and 4) HIV-2 antibody. Specimens that exhibited any level of reactivity were confirmed using the Geenius™ Supplemental HIV 1/2 Antibody assay. RNA testing had to be referred to a different laboratory and therefore this process was reserved for specimens testing positive for HIV p24 antigen without evidence of antibody presence as well as settling any discrepancies between the BioPlex® HIV Ag-Ab assay and the Geenius™ Supplemental HIV 1/2 Antibody assay.

2.4 Results

The majority of specimens tested belonged to patients aged 19 to 45 years old (75.6%). Representative of typical low HIV prevalence in Alaska, only 35 specimens demonstrated presence of HIV antibodies and/or p24 antigen (0.28%) during routine HIV screening. Of these

35 positive specimens, 22 (0.18% of total) were confirmed using the Geenius™ assay. The remaining 13 specimens that could not be confirmed were referred to Wadsworth Center, Albany, NY) for nucleic acid testing (NAT) (Supplemental table), with the exception of 3 specimens which did not have enough serum volume left at this stage of the algorithm. Patients were contacted in these cases and specimens were redrawn to provide additional serum to complete the algorithm. One specimen demonstrated positivity for HIV-1 and HIV-2 antibodies, but could only be tested for HIV-1 RNA due to lack of specimen volume. An additional specimen was redrawn in this case and retested to provide serum for HIV-2 RNA NAT. The positive predictive value in the Alaska population tested was 62.9% (TP/(TP+FP), 22/(22+13)). Of the 22 true positives, 21 specimens (95.5%) were positive for HIV-1 antibody only. The other true positive specimen demonstrated positivity for both HIV-1 antibody and HIV-1 p24 antigen. Of the 13 false positive reactions, 5 (38.5%) were positive for HIV-1 antibody only, 5 (38.5%) were positive for HIV-1 p24 antigen only, and 3 (23.1%) were positive for all targets (HIV-1 and HIV-2 antibodies as well as HIV-1 p24 antigen).

Table 2.1 describes average IDX values produced by the BioPlex® HIV Ag-Ab assay for each targeted analyte by result type. Most false positive reactions demonstrated an IDX measurement <4 with the exception of one specimen which measured between 11-12 IDX for HIV-1 antibody. On average, true positive reactions measured 10-fold higher (>180 IDX) when compared to false positive IDX values. Additional details describing test outcomes for specimens testing falsely positive can be found in Table 2.2.

2.5 Discussion

Low prevalence of HIV infection in a population can lend itself to increased false positivity on diagnostic tests. For instance, BioRad Laboratories tested 6,395 patients in a low-risk population and found that 28 were repeatedly reactive for HIV using the BioPlex® HIV assay (specificity of 99% and sensitivity of 100%). Only 19 of these positives could be confirmed with the Geenius™ Supplemental HIV 1/2 assay and/or NAT testing (19/(19+9) = 67.9% PPV) [11], which is similar to this study (62.9%). This demonstrates that the BioPlex assay performed as intended in a low-prevalence setting. Although this isn't optimal, other HIV 1/2 antigen/antibody immunoassays have reported even lower PPVs, such as the Abbott Architect HIV antigen/antibody combination assay which demonstrated a PPV of 31.2% with significant PPV differences when testing sera from males (49.9%) vs. females (2.5%) [6].

Algorithm adjustments have shown to decrease the likelihood of false positives in various populations. In one study, PPV was improved by using two separate HIV 1/2 antigen/antibody immunoassays during the screening process, in this case the Abbott Architect used in conjunction with the Vidas HIV Duo Ultra (97% PPV) [13]. A PPV of 83% was demonstrated on the Abbot Architect by considering significantly lower signal-to-cutoff ratios in a study focused pregnant women, a group that sometimes experiences higher rates of false positive HIV test results [14].

In our study, all false positive results demonstrated low IDX values, with the exception of one specimen. Based on our results, we suggest that caution be applied to any low positive BioPlex® HIV Ag-Ab assay IDX value (< 4) when reporting preliminary HIV results. Also, supplemental antibody testing provided by the Geenius™ assay did not enhance HIV test interpretation. Sensitive RNA testing, however, helped rule out the presence of detectable virus,

which allowed for the final conclusions of each false positive test result. Adjustments to the testing algorithm to bypass the supplemental antibody assay if the BioPlex® HIV Ag-Ab assay generates low positive IDX values may be helpful to reduce cost, labor, and turnaround time by following through directly to NAT testing (Figure 2.1).

The BioPlex® HIV Ag-Ab assay will likely become more common among laboratories conducting HIV surveillance based on its technical ability to further characterize HIV positive specimens as well as reduce labor. In order to address the issue of false positivity in low-prevalence populations, laboratories choosing to adopt the BioPlex® HIV Ag-Ab assay may want to consider bypassing the supplemental confirmation assay for specimens exhibiting low BioPlex® IDX values and proceed to RNA testing. Similar strategies have been considered for BioPlex® HIV-1 p24 Ag positive-only results [15]. This will ensure that weaker reactions obtained during HIV Ag-Ab combination screening are followed up to test for the presence of viral RNA indicating true HIV infection.

2.6 Ethical Approval

This research project was reviewed and exempted by the University of Alaska Fairbanks Institutional Review Board (IRB) (Approval letter No. 1219176-1).

2.7 Acknowledgements

Authors represent the State of Alaska's Government, Department of Health and Social Services, Division of Public Health, Section of Laboratories. The authors wish to thank the staff at the Alaska State Public Health Virology Laboratory, especially Mary Louise Walmsley, for her technical expertise in the evaluation of the laboratory specimens, and Dr. Barbara Werner from the Association of Public Health Laboratories for her assistance with editing the manuscript.

2.8 References

1. Alonso, R., et al., *Evaluation of the Architect HIV Ag/Ab Combo Assay in a low-prevalence setting: The role of samples with a low S/CO ratio*. Journal of Clinical Virology, 2018. **103**: p.43-47.
2. Lang, R., et al., *HIV misdiagnosis: A root cause analysis leading to improvements in HIV diagnosis and patient care*. Journal of Clinical Virology, 2017. **96**: p. 84-88.
3. Hardie, D.R., et al., *Contamination with HIV antibody may be responsible for false positive results in specimens tested on automated platforms running HIV 4th generation assays in a region of high HIV prevalence*. PLOS ONE, 2017. **12**(7): p. e0182167.
4. Chacón, L., M.L. Mateos, and Á. Holguín, *Relevance of cutoff on a 4th generation ELISA performance in the false positive rate during HIV diagnostic in a low HIV prevalence setting*. Journal of Clinical Virology, 2017. **92**: p. 11-13.
5. Alonso, R., et al., *Evaluation of the Architect HIV Ag/Ab Combo Assay in a low-prevalence setting: The role of samples with a low S/CO ratio*. Journal of Clinical Virology, 2018. **103**: p. 43-47.
6. Kim, S., et al., *False-positive rate of a "fourth-generation" HIV antigen/antibody combination assay in an area of low HIV prevalence*. Clinical Vaccine Immunology, 2010. **17**(10): p. 1642-1644.
7. Wesolowski, L.G., et al., *Highlights from the 2016 HIV diagnostics conference: The new landscape of HIV testing in laboratories, public health programs and clinical practice*. Journal of Clinical Virology, 2017. **91**: p. 63-68.
8. Fordan, S., et al., *Comparative performance of the Geenius™ HIV-1/HIV-2 supplemental test in Florida's public health testing population*. Journal of Clinical Virology, 2017. **91**(Supplement C): p. 79-83.
9. Keating, S.M., et al., *Performance of the Bio-Rad Geenius HIV1/2 Supplemental Assay in Detecting "Recent" HIV Infection and Calculating Population Incidence*. Journal of Acquired Immune Deficiency Syndromes, 2016. **73**(5): p. 581-588.
10. Lavoie, S., et al., *Heterophilic interference in specimens yielding false-reactive results on the Abbott 4th generation ARCHITECT HIV Ag/Ab Combo assay*. Journal of Clinical Virology, 2018. **104**: p. 23-28.
11. Bio-Rad Laboratories, BioPlex® 2200 HIV Ag-Ab. 2018. http://www.bio-rad.com/en-us/product/bioplex-2200-hiv-ag-ab?ID=NQD4K915, retrieved January 18, 2019.
12. Eshleman, S.H., et al., *Performance of the BioPlex 2200 HIV Ag-Ab assay for identifying acute HIV infection*. Journal of Clinical Virology, 2018. **99-100**: p. 67-70.

13. Avidor, B., et al., *Evaluation of the virtues and pitfalls in an HIV screening algorithm based on two fourth generation assays - A step towards an improved national algorithm*. Journal of Clinical Virology, 2018. **106**: p. 18-22.

14. Adhikari, E.H., et al., *Diagnostic accuracy of fourth-generation ARCHITECT HIV Ag/Ab Combo assay and utility of signal-to-cutoff ratio to predict false-positive HIV tests in pregnancy*. American Journal of Obstetrics and Gynecology, 2018. **219**(4): p. 408.e1-408.e9.

15. Association of Public Health Laboratories and Centers for Disease Control and Prevention, APHL/CDC HIV demonstration project for HIV Nucleic Acid Testing (NAT) referral. Updated January 2018, https://www.aphl.org/Materials/UPDATED_1-18-WADSWORTH_2017_2018_Submitting_Lab_Instructions.pdf, retrieved January 18, 2019.

Table 2.1: Comparison of BioPlex® HIV ag-ab assay IDX values by result type

Result Type	# specimens	Specimen Transit Time (days)	HIV AgAb	HIV-1 Ab	HIV-1 Ag	HIV-2 Ab
True Positive	22	4.55 ± 1.57	187.49 ± 35.45	187.49 ± 35.45	0.52 ± 0.87	0.17 ± 0.08
False Positive	13	3.85 ± 1.77	2.58 ± 2.90	2.06 ± 3.03	1.09 ± 1.11	0.50 ± 0.60
Negative	12,303	4.52 ± 2.40	<1.00	<1.00	<1.00	<1.00

Table 2.2: Description of false positive results from BioPlex® HIV ag-ab assay

ID	Age	Gender	Risk Factors	BioPlex® HIV Ag-Ab assay Results (IDX values)				Geenius™	NAT
				HIV Ag-Ab	HIV-1 Ab	HIV-1 Ag	HIV-2 Ab		
AK#1	19	Female	Unknown	1.21 (± 0.07)	0.49 (± 0.04)	1.21 (± 0.07)	0.54 (± 0.04)	HIV Neg	HIV1 ND
AK#2	20	Female	Unknown	1.43 (± 0.11)	0.62 (± 0.08)	1.43 (± 0.11)	0.70 (± 0.12)	HIV Neg	HIV1 ND
AK#3	22	Male	Inmate	3.71 (± 0.38)	2.24 (± 0.16)	3.71 (± 0.38)	2.19 (± 0.09)	HIV Neg	HIV1 ND
			redraw	2.14 (± 0.08)	1.51 (± 0.10)	1.91 (± 0.36)	1.57 (± 0.55)	NP	HIV2 ND
AK#4	26	Female	Pregnant	1.07 (± 0.05)	1.07 (± 0.05)	0.17 (± 0.01)	0.08 (± 0.03)	HIV Neg	HIV1 ND
AK#5	35	Female	Unknown	2.77 (± 0.09)	1.01 (± 0.02)	2.77 (± 0.09)	1.07 (± 0.02)	HIV Neg	HIV1 & HIV2 ND
AK#6	38	Male	Unknown	11.77 (± 0.30)	11.77 (± 0.30)	0.08 (± 0.03)	0.06 (± 0.02)	HIV Neg	HIV1 ND
AK#7	39	Female	Pregnant	1.05 (± 0.13)	0.49 (± 0.06)	1.05 (± 0.13)	0.52 (± 0.02)	HIV Neg	QNS
			redraw	0.95	0.48	0.95	0.51	NP	NP
AK#8	45	Male	Inmate	1.01 (± 0.07)	1.01 (± 0.07)	0.24 (± 0.03)	0.18 (± 0.05)	HIV Neg	HIV1 ND
AK#9	67	Male	Unknown	3.19 (± 0.14)	3.19 (± 0.14)	0.11 (± 0.03)	0.08 (± 0.02)	HIV Neg	HIV1 ND
AK#10	40	Female	Pregnant	1.65 (± 0.12)	1.65 (± 0.12)	0.26 (± 0.06)	0.13 (± 0.03)	HIV Neg	QNS
			redraw	1.97 (± 0.08)	1.97 (± 0.08)	0.44 (± 0.10)	0.26 (± 0.03)	HIV Neg	HIV1 ND
AK#11	52	Male	Inmate	1.46 (± 0.32)	0.20 (± 0.03)	1.55 (± 0.20)	0.10 (± 0.02)	NP	HIV1 ND
AK#12	15	Male	Inmate	2.04 (± 0.13)	2.04 (± 0.13)	0.42 (± 0.09)	0.25 (± 0.03)	HIV Neg	HIV1 ND
AK#13	36	Male	Unknown	1.13 (± 0.23)	0.95 (± 0.20)	1.13 (± 0.23)	0.64 (± 0.08)	HIV Neg	QNS
			redraw	0.77	0.63	0.77	0.45	NP	NP

*Index Values ≥ 1.0 are considered reactive. Reactive specimens tested in triplicate. Confirmation was performed on specimens exhibiting reactivity on any bead result. ND – RNA Not Detected, QNS – quantity of serum not sufficient for further evaluation, NP – not performed. The BioPlex® HIV Ag-Ab assay is a screening method while the Geenius and NAT (nucleic acid amplification test) are considered confirmatory methods. Specimens AK#3, #7, #10, and #13 were redrawn and tested again due to insufficient quantities from the initial blood draw for confirmation testing.

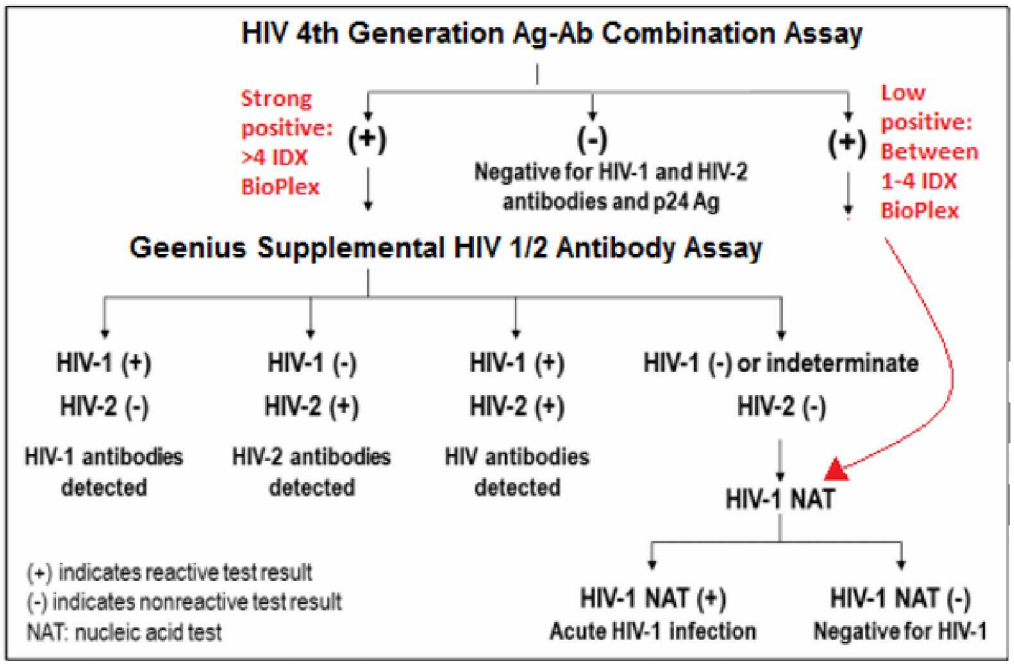


Figure 2.1 Current recommended HIV testing algorithm showing suggested modifications in red [2].

Chapter 3 Application of Next Generation Sequencing for the Detection of Human Viral Pathogens in Clinical Specimens[3]

3.1 Abstract

Next generation sequencing (NGS) is a new technology that can be used for broad detection of pathogens and is rapidly becoming an essential platform in clinical laboratories. It is not known how NGS will displace or enhance gold standard methodologies in infectious disease diagnosis. To investigate the feasibility and application of NGS technology in public health laboratories and compare NGS technology with conventional methods, Illumina's MiSeq system was used to detect viral pathogens alongside other conventional virology methods using typical clinical specimen matrices. Sixteen clinical specimens and two CDC proficiency panels containing seventeen specimens were analyzed. Known pathogenic viral nucleic acid was positively identified in all clinical specimens, correlating and building upon results obtained by more conventional laboratory methods. Sequencing depths ranged from 0.008X to 319X and genome coverage ranged from 0. 6% to 99.9%. To qualify the described methods used to analyze data derived from clinical specimens, the results of a clinical proficiency panel are also presented. These results reveal true scarcity of known pathogenic viral nucleic acids in clinical specimens. NGS outperforms conventional detection methods in this study by turnaround time as well as the improved depth of knowledge in regards to serotyping and drug resistance.

---

[3] This chapter blends and expands upon the following two manuscripts which can be found in their exact published forms in Appendices B and C, respectively:

Parker J, Chen J. *Application of next generation sequencing for the detection of human viral pathogens in clinical specimens*. Journal of Clinical Virology. 2016; **86**:20-26

Parker J, Chen, J. *Next generation sequencing in clinical virology diagnostics*. Clinical Lab International, Feb/Mar 2017:6-9.

3.2 Introduction

Methodologies to detect pathogenic viruses in clinical specimens have transitioned from classic cell culture and antibody-antigen techniques to more sensitive molecular methods such as polymerase chain reaction (PCR). The targeted nature of these methodologies hinders their ability to accommodate the true diversity of human pathogens in a clinical specimen, especially viruses [1]. Next generation sequencing (NGS) technologies are quickly demonstrating their ability to provide broad detection of infectious agents in a target-independent manner [2-7]. NGS has many advantages beyond the improved detection of all suspected, unsuspected, or even novel pathogens in a clinical specimen [8]. Familiarization with pathogen genomic sequences within clinical specimens enhances our understanding of infectious disease through further discovery of pathogen variability and genotyping [9-13] , drug resistance or response to therapy [14-16], vaccine development and efficacy monitoring [17], and further characterization of the metagenome [18, 19]. The use of NGS for routine use in clinical diagnostics is emerging with its own set of limitations and challenges [13, 20]. Focusing on viruses of public health importance, we compared the performance of NGS alongside other more common viral detection methodologies.

3.3 Methods and Materials

3.3.1 Specimens

Sixteen previously tested clinical specimens, swab and serum specimens, were provided by the Alaska State Virology Laboratory in Fairbanks, Alaska. Two proficiency panels with a combined seventeen specimens for detecting antiviral resistance markers in the neuraminidase gene of influenza A virus were also tested as a quality indicator of this

process. Proficiency specimens consisted of cultured Madin-Darby Canine Kidney Epithelial (MDCK) cells infected with influenza A virus.

3.3.2 Construction of sequencing library

Nucleic acid was isolated from 500µL of the original clinical specimen using phenol/chloroform followed by ethanol precipitation. DNA or RNA molecules were selected by using DNase I (serum and proficiency specimens) or RNAse (swab specimens, with the exception of the influenza specimens). Quantity was evaluated using the Qubit instrument (Thermo Fisher Scientific) and the Agilent Bioanalyzer. The Nextera DNA Specimen Preparation protocol (Illumina, [21]) and the NEBNext Ultra RNA Library Prep Kit protocol (New England Biolabs, [22]) were followed to prepare sequencing libraries.

3.3.3 Next generation sequencing and data analysis

Libraries underwent paired-end sequencing on the Illumina MiSeq using a v.2 500-cycle kit. Read files were imported into Pathseq™Virome for reference genome identification. Alignments to the identified viral genome sequence(s) were performed by Sequencher (v5.1) in addition to an external tool, Genomic Short-read Nucleotide Alignment Program (GSNAP) [23]. Read depth and genome coverage was established using Tablet (v.1.13.12.17, [24]).

3.3.4 Conventional Viral Detection Methods

Tests performed in addition to NGS on the swab specimens included cell culture, fluorescent microscopy, and serum neutralization for serotyping isolates that were

successfully grown in cell culture. Tests performed in addition to NGS on the serum specimens included an enzyme-linked immunosorbent assay (ELISA) to determine the presence of HCV antibodies (Roche). Proficiency specimens were also screened by real-time quantitative PCR (qPCR) prior to NGS. Brief descriptions of these more conventional methodologies are described next.

3.3.4.1 Viral Culture

A portion of each specimen (0.3mL) was inoculated into fresh cell monolayers of MRC-5, HEp-2, and RMK cells (Diagnostic Hybrids). Infected cultures were then placed in 35°C (±2°C) incubators. Cultures were observed for cytopathic effect (CPE) by light microscopy periodically over 1-14 days. CPE was rated on a scale of 1+ to 4+, where 1+ =25% cells with CPE, 2+ =50% cells with CPE, 3+ =75% cells with CPE, and 4+ =100% cells with CPE.

3.3.4.2 Fluorescent Microscopy

When CPE was observed (>2+), infected cell monolayers were scraped, dispensed in duplicate onto 8-well slides, dried, and fixed with acetone. Once fixed, slides were dried and subjected to various commercial fluorescent stains containing specific monoclonal or polyclonal antibodies targeting a particular virus. Procedures were carried out according to manufacturer instructions [25, 26] and fluorescence was assessed using a fluorescent microscope and positive (in-house lots) and negative control slides.

3.3.4.3 Serum Neutralization (SN) assay

Once CPE was observed and the virus was identified as an adenovirus, the culture was diluted 1:10 and subjected to incubation in the presence of different monoclonal antibodies representing various serotypes of adenovirus. Initially, each chosen monospecific antisera was heat inactivated at 56°C for 30 minutes. A 1:10 dilution of the isolate was added 1:1 v/v to the heat-inactivated antisera. The antigen-antibody complexes were allowed to react for 1 hour at room temperature. Following incubation, HEp-2 cell monolayers were infected with each antigen-antibody reaction and allowed to incubate for 1.5 hours at 37°C. The growth media was then replaced with cell culture maintenance media. Viable virus was allowed to grow at 33°C on a test tube roller. Cultures were checked for CPE for up to 7 days using light microscopy. Cultures producing no CPE, suggesting the virus was successfully neutralized, were considered to contain that particular serotype of adenovirus.

3.3.4.4 Enzyme-Linked Immunosorbant Assay (ELISA)

Specimens were processed per manufacturer instructions using the EVOLIS analyzer (BioRad, Inc.) in combination with the HCV v.3 antibody detection kit (Roche) [27].

3.3.4.5 Real-time quantitative PCR (qPCR)

qPCR was performed as previously described to screen specimens for influenza prior to preparing the specimens for NGS [28, 29].

3.3.5 Comparative Sequencing Method - Pyrosequencing

Consensus results from pyrosequencing (Roche 454) were obtained from laboratories participating in the proficiency panel for the detection of antiviral-resistant influenza A/H1N1pdm09 and influenza A/H3N2 viruses. These results were compared to those obtained by the next generation sequencing process described in section 3.3.3.

For pyrosequencing, specimens were prepared by first performing PCR to create amplicons of a portion of the neuraminidase gene. Particular amino acids motifs, as translated from bases in this amplicon, indicate resistance or susceptibility to current influenza neuraminidase-inhibitors (oseltamivir, zanamivir, and peramivir). Wild-type influenza A/H1N1pdm09 has two known amino acid positions that lessen their susceptibility to current influenza antivirals: H275Y (histidine to tyrosine) and I223K (isoleucine to lysine). Wild-type influenza A/H3N2 viruses have three known amino acid positions that lessen their susceptibility to current influenza antivirals: E119V (glutamic acid to valine), R292K (arginine to lysine) or N294S (asparagine to serine). Wild-type influenza B viruses have seven known amino acid positions that lessen their susceptibility to current influenza antivirals: E105K (glutamic acid to lysine), E117A/G (glutamic acid to alanine or glycine), Q138K/R (glutamine to lysine or arginine), P139S (proline to serine), G140R (glycine to arginine), D197N/E (aspartic acid to asparagine or glutamic acid), I221T/R/V/L (isoleucine to threonine, arginine, valine, or leucine), H273Y/R (histidine to tyrosine or arginine), and N294S (asparagine to serine).

3.4 Results

3.4.1 NGS for detecting clinical adenovirus infections

Two infections were able to be diagnosed and further characterized using NGS (Figure 3.1). The first, a nasopharyngeal swab from a two-year-old male, was cultured and identified as serotype 2 adenovirus. This process took a total of 14 days to complete. PathSeq Virome analysis identified Human Adenovirus C strain human/ARG/A51932/2002/2 (JX173079) among the NGS read files. Read coverage was low for this specimen (0.4X) covering only 26.7% of the adenovirus genome. This included 30.2% of the hexon gene (3 fragments, average length 292bp), which is used to differentiate adenovirus [30]. The 3 gene fragments sequenced from this region accurately aligned with reference sequence NCBI AC_000007.1 with >99% similarity to adenovirus. Two of the regions further identified the virus as a serotype 2 adenovirus.

The second specimen, a nasal swab from a 29-year-old female, was cultured and demonstrated adenovirus-like CPE at day eight and was successfully typed as serotype 3 adenovirus 18 days later. PathSeq Virome analysis identified Human Adenovirus B strain (human/USA/ UFL_Adv3a51/ 2007/3 (KF268123)). Alignment to this reference strain using GSNAP showed 585 reads with 3.7X depth and 87.9% genome coverage.

3.4.2 NGS for detecting clinical herpesvirus infection

The herpesvirus results of three specimens were further characterized using NGS (Figure 3.2). Clinical diagnostic tests were performed on two swabs taken from genital lesions, one from a 36-year-old female and another from a 41-year-old female. Herpesvirus was suspected in both cases and viral culture was chosen to diagnose or rule-

out the infection. Both specimens demonstrated CPE overnight and the one was clearly diagnosed as a herpesvirus 1 infection by immunofluorescent staining. PathSeq Virome analysis of the NGS data identified human herpesvirus 1 strain H129 (GU734772). When aligned to this reference sequence using GSNAP, a total of 3,664 read files aligned with 5.3X depth and 79.4% genome coverage.

The other specimen was also cultured, but subsequent staining was indeterminate and had to be repeated with various wash buffers to optimize the assay and get a positive identification of a human herpesvirus 2. An aliquot of the culture isolate was extracted and underwent NGS to confirm. PathSeq Virome analysis identified read files belonging to Human herpesvirus 2 strain SD90e (KF781518). When NGS files were aligned to this reference sequence using GSNAP, 171,146 reads were present representing 99.3% of the genome at 199X depth. This example demonstrates the ability of non-targeted NGS to confirm indeterminate findings from traditional, targeted methodologies. As expected, the overall proportion of viral reads is much higher when sequencing clinical isolates (>15%) as compared to raw clinical materials (<1%).

Human herpesvirus 5 (cytomegalovirus) was found to be present in a third specimen, a nasopharyngeal swab taken from a 3-year-old female. As is typical for slow growing cytomegalovirus, the culture incubated for 14 days before showing distinctive CPE for staining. PathSeq Virome analysis indeterminately identified the cytomegalovirus as human herpesvirus 5 strain HAN2 (JX512200). GSNAP alignment to this reference only identified 8 reads (0.008X, 0.6% genome coverage). However, PathSeq Virome analysis more definitively identified torque teno virus isolate US32 (AF122921). A GSNAP alignment to this reference identified 585 read files (3.7X, 88%

genome coverage). Surveillance for torque teno viruses is not common since they are thought to be ubiquitous in humans and lack concrete disease association, unlike cytomegalovirus [31, 32]. Cytomegaloviruses are also ubiquitous in the human population; however, they have direct associations with disease when latent viruses are reactivated. The risk of reactivation is higher for immune suppressed patients which, if not iatrogenic, may be due to underlying medical conditions [32]. The results of this specimen demonstrate the benefit of combining methodologies in some circumstances to get a clearer interpretation of the patient's status. Low sequence coverage of the cytomegalovirus isolate is thought to be due to such low representation overall compared to the background metagenome (0.0005%).

3.4.3 Further characterization of viral hepatitis C and G viruses

Methodologies to compare NGS's ability to detect hepatitis C was not evaluated, but rather sequence analysis was performed to discover how NGS could be used to further characterize the virus. Specimens were all positive for genotype 1a HCV and specific isolates were identified by PathSeq™Virome (Figure 3.3).

New direct acting antiviral therapies have been designed to target and impair the functions of non-structural proteins, NS3/4A, NS5A, and NS5B. In response, breakthrough mutations in these particular viral genes have demonstrated antiviral resistance to certain HCV antiviral therapies [33-37]. Figure 3.3 summarizes resistance data obtained by NGS for each protein targeted as well as the antiviral therapies that are associated. Isolate identified as V60-like was found to have one mutation affecting susceptibility to NS5A-inhibitors (M38V) and 2 mutations affecting susceptibility to

N3/NS4-inhibitors (T54S and Q80K). Three other isolates, identified as V179-, V173-, and V269-like, each showed 1 mutation leading to reduced susceptibility to N3/NS4-inhibitors (Q80K).

Figure 3.3 also shows that hepatitis G virus (HGV) isolates were simultaneously identified in three of the five specimens tested (60%). It is known that HGV infections are closely associated with HCV infections due to parallel routes of transmission [38, 39]. Although not commonly practiced, more thorough surveillance of HGV infections in humans may be necessary due to indications that it is a significant player in determining the course and prognosis of other diseases such as HIV, HCV, and even diseases of the brain [5, 39]. NGS is an appropriate method to detect and characterize both HCV and HGV viruses in parallel.

3.4.4 Antiviral resistance of influenza viruses in clinical specimens

Data is provided for six nasopharyngeal swab specimens containing influenza viruses tested by PCR methods and NGS (Figure 3.4). NGS results detected regions of the genome attributed to antiviral drug resistance, as previously described [40-42]. Genome-wide diagrams of alignments to reference genomes produced by Tablet illustrate the various coverages obtained for each clinical specimen (Figure 3.4). Large variations amongst the results are attributed to the quantity of virus in the original clinical specimen since enrichments techniques such as filtration or centrifugation were not used. Identification of polymorphisms, especially those occurring at specific positions of the neuraminidase gene, may indicate various levels of antiviral resistance to neuraminidase inhibitors and serves as an important piece of information in terms of influenza

surveillance. No variants were detected amongst the clinical specimens analyzed (all wild type). For influenza A/H3, one specimen was missing sequence information for one amino acid motif, 119. For influenza B, one specimen was missing the entire NA gene and could not be analyzed.

Results of a proficiency panel intended for laboratories using pyrosequencing methods are compared to those obtained by NGS (Table 3.1). Results of NGS had 100% concordance with pyrosequencing results for distinguishing wild-type and variant viruses by identifying mutations in the specific amino acid motifs in the neuraminidase gene as an indicator of antiviral resistance. NGS revealed low coverage sequence reads in negative specimens (10 reads each for PT-A-1 and PT-A-4) revealing the need to establish a standard in which to confidently distinguish non-specific and specific reads in NGS data.

3.4.5 Non-specific viral sequencing reads in NGS data

Low coverage viral sequencing is an issue when working with clinical specimens. Collections from patients represent a wide array of pathogen quantities. Although sequencing is becoming more competitive with other conventional methodologies in terms of cost, there is a need for increasing sequencing depth in order to detect pathogenic viruses in clinical specimens. Deeper sequencing also allows for a greater chance of detecting non-specific viral reads. For instance, other viruses were considered "poorly" detected by Pathseq™Virome in each clinical specimen; however, these results could not be substantiated by GSNAP alignments due to very few reads and coverages.

Like other clinical assays, NGS needs a cutoff to determine the true presence of a pathogen versus carry-over or contamination between specimens or other non-specific reads. True negative proficiency specimens (PT-A-1 and PT-A-4) contained 10 reads aligning to the H1N1 NA reference gene (false positive) whereas the human herpesvirus 5 that grew from a nasopharyngeal swab also only had 10 reads. Indeterminate results, such as these, may need to undergo repeat testing where more involved enrichment techniques can be employed to determine the true presence of a virus in low-titer clinical specimens.

3.5 Discussion

Applications of NGS in a clinical laboratory were considered to characterize pathogenic viruses of public health importance. NGS data was compared to data obtained by more conventional methods of virus detection. DNA viruses such as adenoviruses and herpesviruses and RNA viruses such as hepatitis C virus and influenza virus were further characterized in these experiments. In most cases, with the exception of one herpesvirus, information retrieved by NGS met or exceeded that of conventional methodologies. Beyond mere detection, NGS proved to be a laboratory tool capable of predicting the effects of drug treatment as well.

The importance of characterizing pathogenic viruses in clinical specimens should not be undermined by the fact that very few treatments are available for viral infections. However, for infections caused by viruses with available antiviral drug treatments, known resistance markers in the viral genome could be routinely monitored using NGS. This would allow for surveillance of drug effectiveness against circulating strains of viruses in particular populations.

Adenoviruses are widely distributed amongst humans and generally cause only mild acute respiratory illness. However, some serotypes are more virulent and can be associated with outbreaks, severe pneumonia, and possibly cancer such as serotypes 14, 55, and 12 [30, 43]. As recent as November 2018, serotype 7 was implicated in an outbreak of respiratory illness in New Jersey where 11 children died [44]. NGS, which was complete in approximately 4 days compared to the 14 – 18 days required for culture and serotyping, demonstrated to be a powerful tool for characterizing adenovirus infection when compared to traditional methodologies. Routinely sequencing adenoviruses could prove useful in monitoring how closely circulating strains match vaccine candidates for serotypes 4 and 7, two viruses responsible for severe respiratory illness among military recruits [45].

Other DNA viruses, such as herpes simplex viruses, grow very quickly in culture, but particular strains can be difficult to positively identify using immunofluorescence assays alone due to non-specific staining. NGS confirmed indeterminate staining results and further characterized the virus isolate as a specific human herpesvirus 2 strain. Sequencing data with this level of resolution of the herpesvirus genome could be used to identify transmission routes to relate infections between people. This type of information could be used to determine the source of the infection as well as bring additional benefit to epidemiologic and forensic data. Like adenovirus infections, herpes simplex viruses have available antiviral treatment options. Sequencing data could be used to evaluate the effectiveness of antivirals on specific strains by routinely monitoring markers of antiviral resistance.

One herpesvirus infection caused by cytomegalovirus (HHV5) was unable to be adequately sequenced and characterized due to suspected low initial viral titer in the original clinical specimen (i.e. minimal starting template). Culturing the specimen, in this case, could

amplify the virus into higher titers that could provide more starting template to be sequenced, much like that which was demonstrated in the confirmation of the human herpesvirus 2 specimen. Some hesitancy to culture the virus is understandable, especially when working with viruses with high mutation rates, like RNA viruses. DNA viruses, like herpesviruses, have lower rates of mutation in comparison to their single-stranded counterparts. Therefore, culturing has less of an effect on the final sequences of what is considered to represent the original herpesvirus genome.

Hepatitis C virus (HCV) is a growing concern for public health and tends to be difficult to design targeted methodologies around due to the high variability of viral genomes known, even within the same patient. NGS is a powerful tool for characterizing HCV infections and, from this experience, more informational than targeted genotyping assays. Since these methods produce sequences of the entire HCV genome (coverage ranged from 92.4% - 95.6%), data could be generated describing the mutations at key locations across the genomes that are known to cause drug resistance.

Antiviral resistance is also critical when characterizing current circulating influenza virus strains and NGS was able to identify viruses that would be considered susceptible to neuraminidase inhibitors. In two cases, the viral load of the specimen was too low to achieve good genome coverage across the neuraminidase gene, but this issue could be resolved by screening specimens for high titer (i.e. qPCR) or utilizing enrichment techniques such as ultracentrifugation, filtration of other background nucleic acid, or viral isolation via culture.

Through increased use of NGS technologies, reference databases of whole genome sequences of viral pathogens can grow and enhance our ability to more definitively identify

sequencing reads. Although these experiments describe concurrency between conventional methods and NGS, there was also evidence of unexpected viruses, such as of the presence of co-infecting viruses like hepatitis G and torque teno virus. The standard 4-day turnaround time needed to complete NGS could be improved with extraction and library preparation automation, as well as advances in sequencing technology (each run ~40 hours). Based on these outcomes and the growing body of literature, NGS will change our approach as clinical laboratorians and improve our ability to detect and more fully characterize agents of infectious disease in clinical specimens in a non-targeted manner. These improvements to pathogen characterization can also be a powerful epidemiological tool in its ability to relate cases, as we will see in Chapter 4.

3.6 Competing Interests

The authors declare no competing interests or conflict of interest.

3.7 Ethical Approval

This research project was reviewed and approved by the University of Alaska Fairbanks Institutional Review Board (IRB) (Approval letter No. 667418-2).

3.8 Acknowledgements

Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103395 and by an equipment grant of Alaska State Public Health Laboratories. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the NIH and Alaska State Public Health Laboratories. All authors are employees of the State of Alaska's Government, Department of Health and Social

Services, Division of Public Health, Section Laboratories. The authors wish to thank the staff at the Alaska State Public Health Virology Laboratory for their technical expertise in the evaluation of the laboratory specimens.

Table 3.1: NGS proficiency compared to pyrosequencing methodology for detecting antiviral resistance in influenza A virus

			Pyrosequencing			NGS					
			275		223	275		223			
H1N1	Real-time PCR	Variant	H%	Y%	I%	H%	Y%	I%	NA Reads	% NA	Depth
PT-A-1	ND	indeterminate	--	--	--	--	--	--	10	35.9%	1.3
PT-A-2	H1N1pdm09	H275	100	0	100	100	0	100	5728	99.6%	850
PT-A-3	H1N1pdm09	H275Y	<10	>90	100	8	92	100	275	98.6%	38.3
PT-A-4	ND	indeterminate							10	48.0%	1.6
PT-A-5	H1N1pdm09	H275Y	60-65	35-40	100	68	42	100	3165	99.8%	436
PT-A-6	H1N1pdm09	H275	100	0	100	100	0	100	2589	98.9%	361
PT-A-7	H1N1pdm09	H275	100	0	100	100	0	100	1364	99.1%	204
PT-A-8	H1N1pdm09	H275Y	60-65	35-40	100	59	41	100	3624	99.7%	497
PT-A-9	H1N1pdm09	H275Y	100	0	100	100	0	100	8878	99.6%	1427
PT-A-10	H1N1pdm09	H275, I223K	100	0	0	100	0	0	4924	99.6%	764
PT-B-1	Not tested	H275	100	0	100	100	0	100	12,788	99.8%	639
PT-B-2	Not tested	H275Y	0	100	100	0	100	100	8,031	99.9%	402
PT-B-3	Not tested	H275, I223K	100	0	0	100	0	0	10,139	99.9%	508
PT-B-4	Not tested	H275Y	60-65	35-40	100	60	40	100	11,496	99.9%	575
H3N2			E119	R292	N294	E119	R292	N294			
PT-B-5	Not tested		100	100	100	100	100	100	8,059	99.9%	418
PT-B-6	Not tested	E119V	0 (V100%)	100	100	0 (V100%)	100	100	12,878	99.9%	670
PT-B-7	Not tested	R292K	100	0 (K100%)	100	100	0 (K100%)	100	30,550	99.9%	1,585

Two separate panels of proficiency specimens (A & B) were tested using NGS and compared to results obtained by laboratories performing pyrosequencing for currently circulating influenza A viruses. Known amino acid positions are unique to each type of neuraminidase gene (N1 or N2) separating out the viruses and varying the positions monitored for antiviral resistance.

Specimen Type	Origination	Culture	Staining	Serum Neutralization Assay	NGS
Nasopharyngeal Swab	2yo M	Hep#2[a] CPE at Day 2 (3+)[b]	DFA[c], Adenovirus	Adenovirus (type 2) TAT: 14 days	Human adenovirus C strain human/ARG/A15932/2002/2 TAT: 4 days
Nasal Swab	29yo F	Hep#2[a] CPE at Day 8 (4+)[b]	DFA[c], Adenovirus	Adenovirus (type 3) TAT: 18 days	Human adenovirus B strain human/USA/UFL_Adv3a51/2007/3 TAT: 4 days

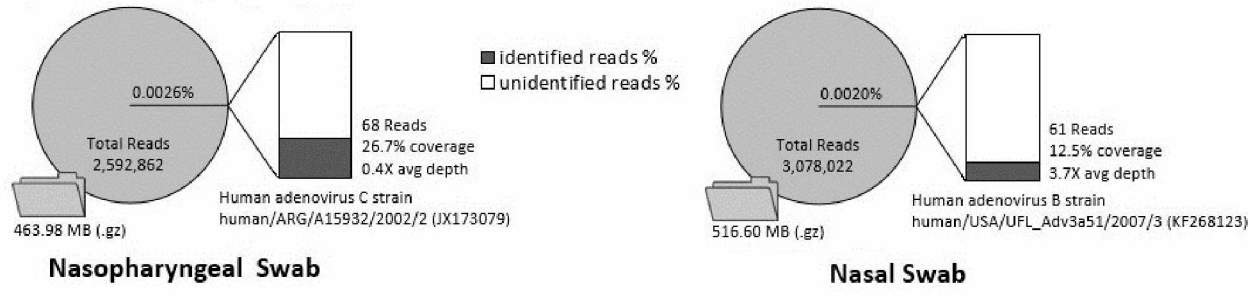


Figure 3.1: Comparison of NGS and conventional virology assays for detecting adenovirus infection. Hep#2 cells, Diagnostic Hybrids; CPE, cytopathic effect, CPE scale, 1+ (25% cells with CPE), 2+ (50% cells with CPE), 3+ (75% cells with CPE) and 4+ (100% cells with CPE); DFA, Direct Fluorescence Assay, Diagnostic Hybrids; TAT, turnaround time indicating days needed to complete the analysis process. Pie charts show the total number of reads (and data storage needed in compressed files, .gz) in relation to the number of viral reads obtained for a particular virus. Exploding bar graph shows the percentage of the full genome that was identified by the reported number of reads.

Specimen Type	Patient	Culture	Staining	NGS → TAT: 4 days
Genital Lesion Swab	36yo F	MRC-5 3+ CPE - Day 1	DFA, Human herpesvirus 1 →TAT: 1 day	Human herpesvirus 1 strain H129
Genital Lesion Swab	41yo F	MRC-5 3+ CPE - Day 1	DFA, Human herpesvirus 2 →TAT: 1 day	Human herpesvirus 2 strain SD90e
Nasopharyngeal Swab	3yo F	MRC-5 1+ CPE at Day 14	IFA, Human herpesvirus 5 →TAT: 14 days	Torque Teno virus isolate US32 & Human herpesvirus 5 (trace)

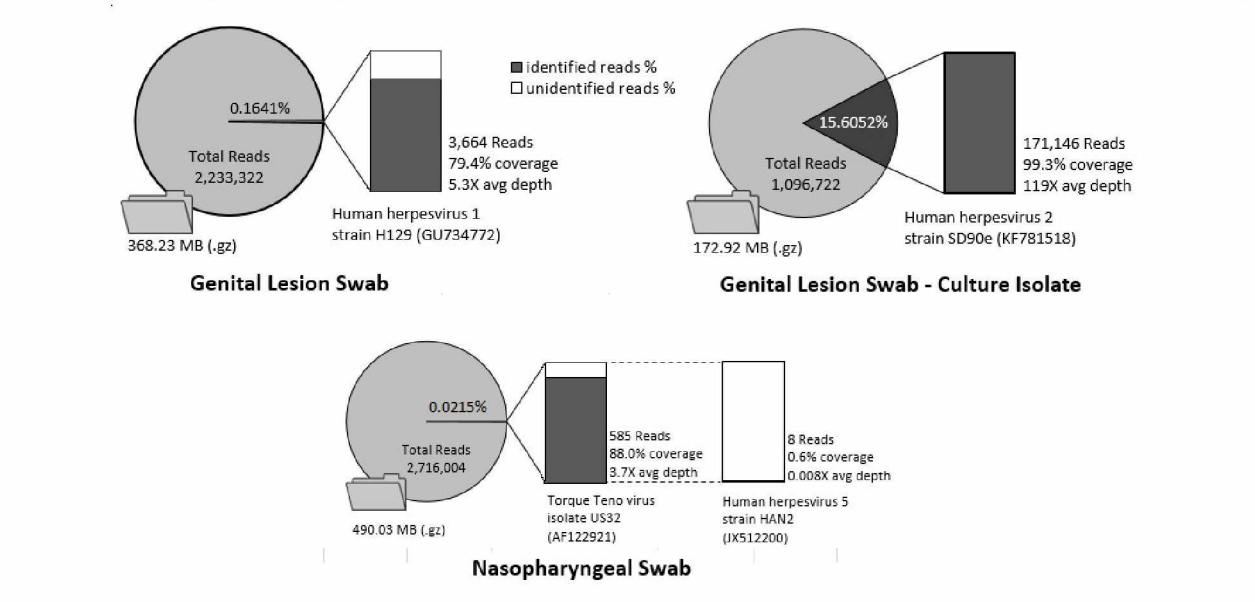


Figure 3.2: Comparison of NGS and conventional virology assays for detecting herpesvirus infection. Three clinical specimens from patients infected with different types of herpesvirus were tested by NGS and the results were compared with those obtained by different methodologies. MRC-5 cells, Diagnostic Hybrids; IFA, immunofluorescence assay, Bartel's reagent.

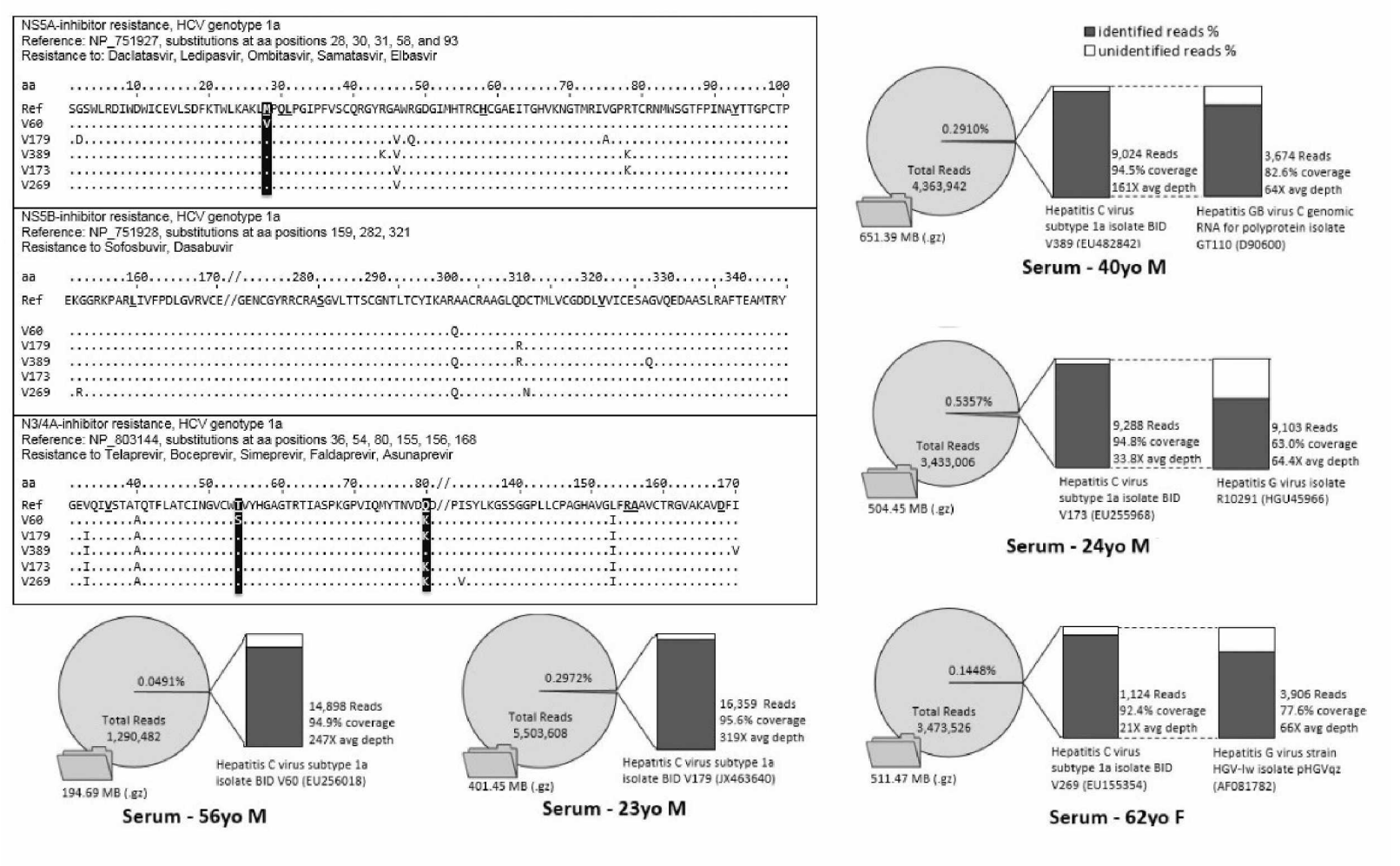


Figure 3.3: Detection and characterization of hepatitis C and G viruses in 5 different sera using NGS.
Three groups of antiviral drugs are shown targeting proteins NS5A, NS5B, and N3/N4A. Isolates (i.e. V60, V179) should be thought of as V60-like or V179-like, as identified by sequence analysis. Blackened columns of amino acids indicate that one or more isolates were found to have mutations at those positions associated with antiviral resistance.

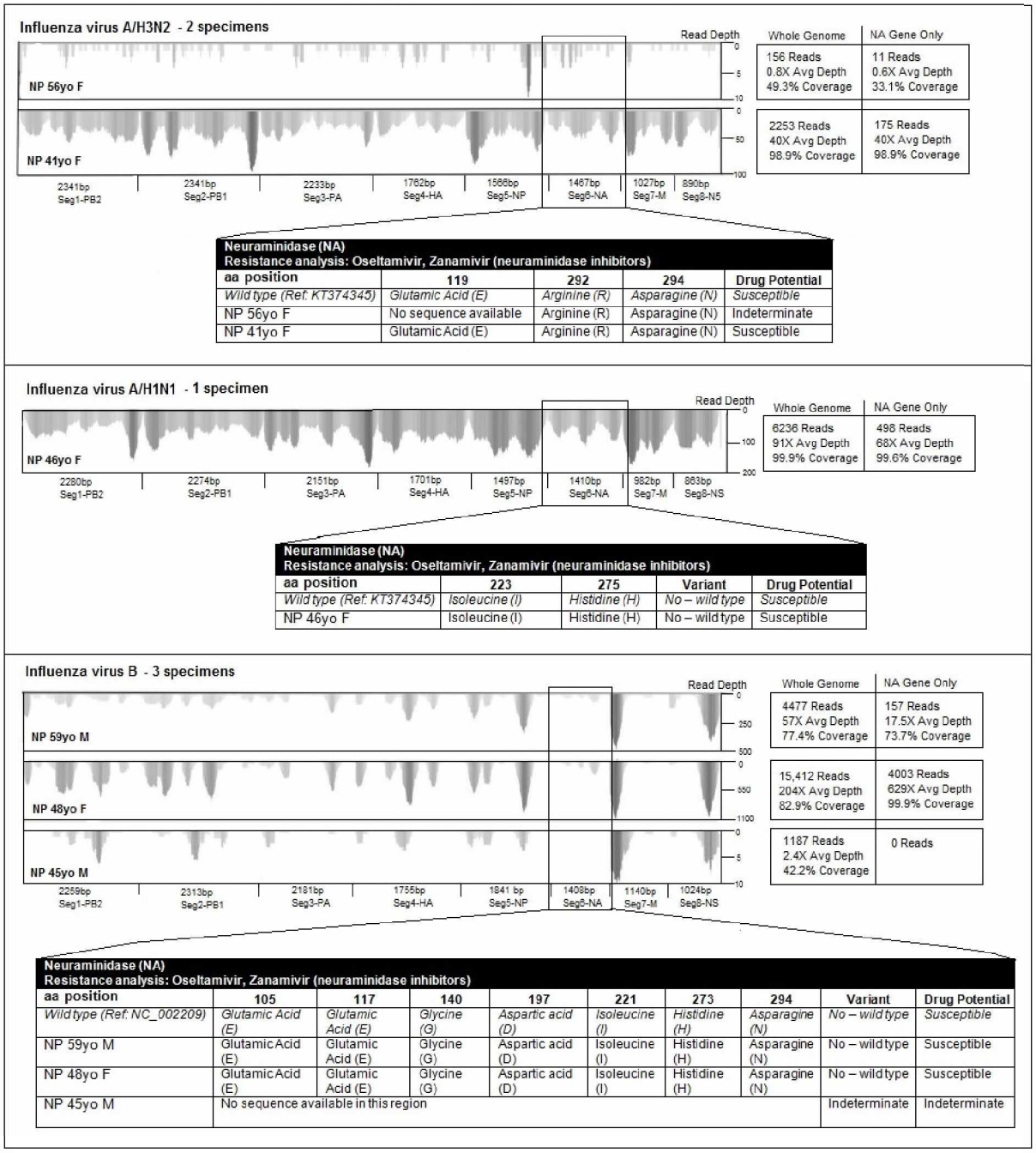


Figure 3.4: Antiviral resistance characterization of influenza viruses using NGS. Genome-wide views as produced by Tablet. Table is split into 3 influenza virus types, A/H3N2 viruses, A/H1N1 viruses, and B viruses. Whole genome metrics are compared to those obtained for the NA gene only. Segment 6, the NA gene, is boxed out for each specimen and exploded into a table describing the outcomes of antiviral resistance analysis performed for each type of influenza virus.

3.9 References

1. Köser, C.U., et al., *Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology.* PLOS Pathogens, 2012. **8**(8): p. e1002824.

2. Bzhalava, D., et al., *Unbiased Approach for Virus Detection in Skin Lesions.* PLOS ONE, 2013. **8**(6): p. e65953.

3. Cheval, J., et al., *Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples.* J Clin Microbiol, 2011. **49**(9): p. 3268-75.

4. Chan, B.K., et al., *Deep Sequencing to Identify the Causes of Viral Encephalitis.* PLOS ONE, 2014. **9**(4): p. e93993.

5. Kriesel, J.D., et al., *Deep Sequencing for the Detection of Virus-Like Sequences in the Brains of Patients with Multiple Sclerosis: Detection of GBV-C in Human Brain.* PLOS ONE, 2012. **7**(3): p. e31886.

6. Moore, R.A., et al., *The Sensitivity of Massively Parallel Sequencing for Detecting Candidate Infectious Agents Associated with Human Tissue.* PLOS ONE, 2011. **6**(5): p. e19838.

7. Yozwiak, N.L., et al., *Virus Identification in Unknown Tropical Febrile Illness Cases Using Deep Sequencing.* PLoS Neglected Tropical Diseases, 2012. **6**(2): p. e1485.

8. Radford, A.D., et al., *Application of next-generation sequencing technologies in virology.* J Gen Virol, 2012. **93**(Pt 9): p. 1853-68.

9. Arroyo, L.S., et al., *Next generation sequencing for human papillomavirus genotyping.* J Clin Virol, 2013. **58**(2): p. 437-42.

10. Flaherty, P., et al., *Ultrasensitive detection of rare mutations using next-generation targeted resequencing.* Nucleic Acids Res, 2012. **40**(1): p. e2.

11. Meiring, T.L., et al., *Next-generation sequencing of cervical DNA detects human papillomavirus types not detected by commercial kits.* Virology Journal, 2012. **9**(1): p. 164.

12. Sijmons, S., M. Van Ranst, and P. Maes, *Genomic and Functional Characteristics of Human Cytomegalovirus Revealed by Next-Generation Sequencing.* Viruses, 2014. **6**(3): p. 1049-1072.

13. Watson, S.J., et al., *Viral population analysis and minority-variant detection using short read next-generation sequencing.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1614): p. 20120205.

14. Han, Y., et al., *Analysis of hepatitis B virus genotyping and drug resistance gene mutations based on massively parallel sequencing.* J Virol Methods, 2013. **193**(2): p. 341-7.

15. Messiaen, P., et al., *Ultra-deep sequencing of HIV-1 reverse transcriptase before start of an NNRTI-based regimen in treatment-naive patients.* Virology, 2012. **426**(1): p. 7-11.

16. Sahoo, M.K., et al., *Detection of cytomegalovirus drug resistance mutations by next-generation sequencing.* J Clin Microbiol, 2013. **51**(11): p. 3700-10.

17. Mathonet, P. and C.G. Ullman, *The Application of Next Generation Sequencing to the Understanding of Antibody Repertoires.* Frontiers in Immunology, 2013. **4**: p. 265.

18. Lecuit, M. and M. Eloit, *The human virome: new tools and concepts.* Trends in Microbiology. **21**(10): p. 510-515.

19. Yang, J., et al., *Unbiased Parallel Detection of Viral Pathogens in Clinical Samples by Use of a Metagenomic Approach.* Journal of Clinical Microbiology, 2011. **49**(10): p. 3463-3469.

20. Isakov, O., S. Modai, and N. Shomron, *Pathogen detection using short-RNA deep sequencing subtraction and assembly.* Bioinformatics, 2011. **27**(15): p. 2027-30.

21. Illumina, *Nextera DNA Specimen Preparation protocol.* **https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_nextera/nexteradna/nextera-dna-library-prep-reference-guide-15027987-01.pdf**.

22. New England BioLabs, *NEBNext Ultra RNA Library Prep Kit for Illumina.* **https://www.neb.com/~/media/Catalog/All-Products/5601629C7791467AB43AAE7A5B361C21/Datacards%20or%20Manuals/E7530_man.pdf**.

23. Wu, T.D. and S. Nacu, *Fast and SNP-tolerant detection of complex variants and splicing in short reads.* Bioinformatics, 2010. **26**(7): p. 873-81.

24. Milne, I., et al., *Tablet—next generation sequence assembly visualization.* Bioinformatics, 2010. **26**(3): p. 401-402.

25. Trinity Biotech, *Bartels Cytomegalovirus (CMV) Fluorescent Antibody Test V1029-81-29 Rev C.* **http://documents.trinitybiotech.com/Product%20Documents/B1029-81-29C%20EN.pdf**.

26. BioRad Laboratories, *Pathfinder HSV 1 and 2 DFA.* **http://www.bio-rad.com/en-ye/sku/25215-pathfinder-hsv-types-1-2**.

27. Roche, *Hepatitis C Virus Encoded Antigen (Recombinant c22-3, c200 and NS5) ORTHO HCV Version 3.0 ELISA Test System.* March 2009: p. **https://www.fda.gov/downloads/BiologicsBloodVaccines/.../UCM176421.pdf**.

28. Weinberg, G.A., et al., *Field evaluation of TaqMan Array Card (TAC) for the simultaneous detection of multiple respiratory viruses in children with acute respiratory infection.* J Clin Virol, 2013. **57**(3): p. 254-60.

29. Wangchuk, S., et al., *Influenza surveillance from November 2008 to 2011; including pandemic influenza A(H1N1)pdm09 in Bhutan.* Influenza Other Respir Viruses, 2013. **7**(3): p. 426-30.

30. Ebner, K., W. Pinsker, and T. Lion, *Comparative sequence analysis of the hexon gene in the entire spectrum of human adenovirus serotypes: phylogenetic, taxonomic, and clinical implications.* J Virol, 2005. **79**(20): p. 12635-42.

31. Kincaid, R.P., et al., *A Human Torque Teno Virus Encodes a MicroRNA That Inhibits Interferon Signaling.* PLOS Pathogens, 2013. **9**(12): p. e1003818.

32. Navarro, D., *Expanding role of cytomegalovirus as a human pathogen.* J Med Virol, 2016. **88**(7): p. 1103-12.

33. Ahmed, A. and D.J. Felmlee, *Mechanisms of Hepatitis C Viral Resistance to Direct Acting Antivirals.* Viruses, 2015. **7**(12): p. 6716-6729.

34. Fridell, R.A., et al., *Genotypic and phenotypic analysis of variants resistant to hepatitis C virus nonstructural protein 5A replication complex inhibitor BMS-790052 in humans: in vitro and in vivo correlations.* Hepatology, 2011. **54**(6): p. 1924-35.

35. Bagaglio, S., et al., *Frequency of Natural Resistance within NS5a Replication Complex Domain in Hepatitis C Genotypes 1a, 1b: Possible Implication of Subtype-Specific Resistance Selection in Multiple Direct Acting Antivirals Drugs Combination Treatment.* Viruses, 2016. **8**(4): p. 91.

36. Lemm, J.A., et al., *Identification of Hepatitis C Virus NS5A Inhibitors.* Journal of Virology, 2010. **84**(1): p. 482-491.

37. Ji, H., et al., *Next generation sequencing of the hepatitis C virus NS5B gene reveals potential novel S282 drug resistance mutations.* Virology, 2015. **477**: p. 1-9.

38. Tanaka, E., et al., *Past and Present Hepatitis G Virus Infections in Areas Where Hepatitis C is Highly Endemic and Those Where It Is Not Endemic.* Journal of Clinical Microbiology, 1998. **36**(1): p. 110-114.

39. Reshetnyak, V.I., T.I. Karlovich, and L.U. Ilchenko, *Hepatitis G virus.* World J Gastroenterol, 2008. **14**(30): p. 4725-34.

40. Farrukee, R., et al., *Influenza viruses with B/Yamagata- and B/Victoria-like neuraminidases are differentially affected by mutations that alter antiviral susceptibility.* J Antimicrob Chemother, 2015. **70**(7): p. 2004-12.

41. Gubareva, L.V., *Molecular mechanisms of influenza virus resistance to neuraminidase inhibitors.* Virus Res, 2004. **103**(1-2): p. 199-203.

42. Ferraris, O. and B. Lina, *Mutations of neuraminidase implicated in neuraminidase inhibitors resistance.* J Clin Virol, 2008. **41**(1): p. 13-9.

43. Doerfler, W., *Adenoviruses.* Medical Microbiology, 1996.

44. State of New Jersey Department of Health, *NJDOH confirms 11th Pediatric Death Of Wanaque Patient; Commissioner Requests MRC Volunteers to Speed Up Separation of Patients.* Retrieved from **https://nj.gov/health/news/2018/approved/20181116a.shtml**, n.d.

45. Potter RN, C.J., Mallak CT, et al. , *Adenovirus-associated Deaths in US Military during Postvaccination Period, 1999–2010.* Emerging Infectious Diseases, 2012. **18**(3): p. 507-509.

Chapter 4 Investigation of a Canine Parvovirus Outbreak using Next Generation Sequencing[4]

4.1 Abstract

Canine parvovirus (CPV) outbreaks can have a devastating effect in communities with dense dog populations. The interior region of Alaska experienced a CPV outbreak in the winter of 2016 leading to further investigation of the virus due to reports of increased morbidity and mortality occurring at dog mushing kennels in the area. Twelve rectal-swab specimens from dogs displaying clinical signs consistent with parvoviral-associated disease were processed using next-generation sequencing (NGS) methodologies by targeting RNA transcripts, and therefore detecting only replicating virus. All twelve specimens demonstrated the presence of the CPV transcriptome, with read depths ranging from 2.2X – 12,381X, genome coverage ranging from 44.8% - 96.5%, and representation of CPV sequencing reads to those of the metagenome background ranging from 0.0015% - 6.7%. Using the data generated by NGS, the presence of newly evolved, yet known, strains of both CPV-2a and CPV-2b were identified and grouped geographically. Deep-sequencing data provided additional diagnostic information in terms of investigating novel CPV in this outbreak. NGS data in addition to limited serological data provided strong diagnostic evidence that this outbreak most likely arose from unvaccinated or under-vaccinated canines, not from a novel CPV strain incapable of being neutralized by current vaccination efforts.

[4] Parker, J., M. Murphy, K. Hueffer and J. Chen. *Investigation of a Canine Parvovirus Outbreak using Next Generation Sequencing.* Scientific Reports, 2017. 7(1): 9633-9633.

4.2 Introduction

Canine parvovirus type 2 (CPV-2) is a non-enveloped, single-stranded DNA virus that causes fatal gastroenteritis in young dogs [1]. The CPV-2 genome is 5323 nucleotides long and possesses at least 2 major open reading frames (ORFs) [2]. CPV-2 infections are associated with significant morbidity and mortality which can reach 91% in untreated pups [3]. Three variants of CPV type 2 are known, CPV-2a, CPV-2b, and CPV-2c, and are highly contagious due to suspected low infective dose requirements combined with high titers of transmissible virus in stools of affected dogs [4, 5]. The virus is highly resistant to environmental conditions and can remain viable outside of its host for at least a year [6].

CPV-2 infections are one of the most common causes of disease outbreaks in dense canine environments such as kennels or shelters, and timely diagnosis is important in order to control the number of affected individuals [7]. CPV is commonly diagnosed in veterinary clinics using rapid fecal enzyme-linked immunosorbent assays (ELISA) that target viral antigen. These tests have high specificity but poor sensitivity when compared to PCR or immune-electron microscopy [8]. Although PCR assays are more sensitive, they can cause difficulty in terms of result interpretation since they can detect live attenuated vaccine strains or produce positive results from dogs showing no symptoms of gastroenteritis. This requires veterinarians to associate PCR results with other clinical signs of CPV infection, the animal's history, and other laboratory parameters such as leukopenia [8, 9]. The performance of any of these antigen-targeting methods are highly variable due to the known phenomenon of intermittent shedding of CPV during the earlier and later stages of disease [10]. Despite these diagnostic challenges, it has been shown that current ELISA and PCR methods are capable of detecting all three variants of CPV in spite of antigenic differences [11].

Between January and April 2016, the interior of Alaska experienced an increased number of CPV cases [12-14]. Outbreaks of canine infectious disease in Alaska can be socially and economically detrimental due to the dense dog population needed to support the state sport, dog mushing. The interior of Alaska, in particular, has a concentrated population of dogs due to the increased presence of professional and recreational mushers operating various sized kennels of 10-100+ sled dogs each. Every winter Alaska hosts many visiting mushers, national and international, who come to the area to train and race their own dogs. Additionally, Alaska hosts several high-profile international dog mushing races, including the Iditarod and the Yukon Quest. These large races frequently include groups of dogs ranging from 350-1350 animals in number, who utilize the same trails, rests stops and parking areas, allowing for extensive comingling and a high potential for disease transmission. Sick dogs can spread the virus through defecation on common trails, where other teams run and transport the virus back to their own kennels. In 2016, the interior of Alaska experienced a mild winter with less than the usual amount of snowfall, leading to accumulation of uncovered fecal material on common mushing trails. In response to the CPV-associated outbreak of disease, mushers were provided notifications throughout the 2016 race season, and were asked to not bring potentially infected dogs to races in order to help slow the progression of the outbreak. Recommendations were also made to isolate sick animals in individual kennels. Due to the perceived increased virulence of the CPV strain or strains associated with the outbreak, additional testing to further characterize the virus was pursued.

Sequence analysis of the VP2 gene, the most abundant and immunogenic protein produced for construction of the viral particle capsid, is used to help subtype and further characterize wildtype CPV [15-18]. Surveillance of this particular protein is critical for assessing

the potential efficacy of the current vaccination strategy and can also be used to relate individual infections in outbreak situations [19-22]. Beyond VP2 investigation, deep-sequencing of the whole genome has been shown to be useful in order to better understand the true nature of CPV molecular diversity and discover new variants [23, 24]. In this experiment, next-generation sequencing (NGS) was used to detect and characterize actively replicating CPV in rectal swabs of canines associated with a suspected outbreak in the interior of Alaska between January - April 2016 by targeting RNA transcripts.

4.3 Materials and Methods

4.3.1 Specimens

Twelve rectal swab specimens representing two communities, A (n=5) & B (n=7), were collected for the sole purpose of disease diagnosis. Communities involved are 258 kilometers (approximately 160 miles) apart. The University of Alaska Fairbanks Institutional Animal Care and Use Committee (IACUC) has determined that this research project did not require IACUC review. Protocol review is not required for diagnostics performed during the course of a disease investigation. All methods were carried out in accordance with relevant guidelines and regulations.

4.3.2 Referral Testing (serology, PCR, & genotyping)

Specimens were sent to Cornell University, Ithaca, NY for evaluation of CPV IgG and IgM antibodies using hemagglutination inhibition (HAI), nucleic acid using PCR, and genotyping (Table 4.1). Total antibody was evaluated, as well as the IgM to IgG ratio upon application of 2-mercaptoethanol to dissociate IgM antibody molecules. Laboratory interpretation guidelines suggest that a 4-fold or greater decrease in titer after 2-mercaptoethanol treatment is evidence of recent parvovirus exposure. Post-vaccination

levels of total antibody can range from 80 to 2,560 HAI, with ranges around 80 demonstrating need for booster vaccination. In addition to serotyping, canine parvovirus PCR was used to rule-in specimens containing CPV nucleic acid and two specimens were genotyped.

4.3.3 Nucleic acid isolation in preparation for sequencing

Flocked swabs were used to collect specimens from the rectum of expired canines presumed to be related to the outbreak (n=12). Confirmation of parvoviral infection was made by rapid fecal enzyme-linked immunosorbent assays (ELISA), necropsy findings consistent with acute parvoviral illness (hemorrhagic enteritis with fibrinous serositis), or both. Swabs were immediately placed in viral transport medium to stabilize viral particles. Representative aliquots (~1mL) of each specimen were centrifuged for 10 minutes at 15,000 rpm to pellet and discard debris as well as the majority of the host and bacterial cells. The supernatant was transferred into a new centrifuge tube and a portion of the supernatant (~500uL) was used as the starting material for nucleic acid isolation using phenol/chloroform followed by ethanol precipitation as previously described [25, 26]. DNAse I (New England Biolabs, Inc.) was added to the isolated total nucleic acid to remove all genomic DNAs.

4.3.4 Library preparation for sequencing

NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Inc) was used to construct sequencing library from a starting quantity of 10-200ng of total RNA. Individual indexes were used to barcode the fragments and allow for specimen pooling. Fragment sizes of ~300bp and larger were selected during AMPure bead cleanup.

4.3.5 Sequencing and analysis

Libraries were pooled and sequenced using the Illumina MiSeq system and the MiSeq Reagent Kit v2 (Illumina) 500-cycle sequencing kit. Paired-end sequencing was performed (2 x 251bp) to achieve all base reads available in the 500-cycle sequencing format. Read files generated by the sequencer were analyzed by GSNAP reference sequence alignment using NCBI Accession NC_001539.1 as the reference genome for parvovirus.

4.4 Results

Serological data was difficult to obtain since affected animals would expire before blood draws could be acquired. Only three dogs in this data set were tested for canine parvovirus antibodies (Parvo6, 7, and 12) each demonstrating low to medium levels of acceptable antibody (Table 4.1). PCR was performed on most specimens in this dataset unanimously ruling-in the presence of CPV nucleic acid. Three specimens were genotyped, Parvo1 and Parvo2 (representing the same animal and therefore an opportunity to assess diagnostic reproducibility) were typed as a 2/2a virus, and Parvo10 as a 2b virus.

Table 4.2 describes the general sequencing metrics for each specimen tested. The number of reads aligning to the canine parvovirus reference sequence (NC_001539.1) ranged from 47 reads for Parvo11 to 263,625 reads for Parvo9. This type of range can be expected when blindly testing clinical specimens and is dependent on sample collection and handling as well as the stage of disease presenting in the canine at the time of specimen collection. Read depth ranged from 2.2X to 12,381X and genome coverage ranged from 44.8% to 96.5% when aligned to the reference genome. Canine parvovirus sequences were heavily masked amongst other sequence

data representing an average of 0.95% (range 0.0015% to 6.7%) of the metagenome across all specimens sequenced.

Whole genome phylogenetic analysis was performed using CLC Workbench 8 (Figure 4.1). Parvo1 and Parvo11 samples were not included in this analysis due their lower than optimum read depths (<30X). Whole genome sequences from the remaining 10 specimens were compared to 13 reference genomes from NCBI representing many variations of subtype 2a and 2b canine parvoviruses as well as two versions of the attenuated vaccine strain. Phylogenetic analysis shows that there were two distinct groups of viruses circulating in the outbreak and these groups were directly associated with the geographical location of the animal when it became ill. Parvo2, 3, 6, &7 formed a group of somewhat older origin and represented animals from Community A. The other grouping of viruses sequenced (Parvo4, 5, 8, 9, 10, & 12) shows more recent evolution and were collected from animals approximately 260 kilometers away in Community B. Interestingly, Parvo6 was a dog that had been transferred from Community B to Community A; however, the sequence data would suggest that the parvovirus infection was actually caused by a virus picked up in Community A, not brought in from Community B.

Further analysis of the VP2 protein of each one of these groupings was performed (Figure 4.2). Parvo6 represented the Community A grouping and Parvo9 represented the Community B grouping due to their high genome coverage percentages and read depths (>30X). The nucleic acid VP2 sequences for these two viruses were 99.4% similar differing by 10 bases. Five (50%) of the base differences resulted in four amino acid sequence differences in the VP2 protein at residues 267, 324, 426, and 440. The amino acid difference at residue 426 is a common codon affected by polymorphism and is used to diagnostically differentiate CPV-2a and CPV-2b viruses with the use of specific sequence-based probes. The amino acid difference at position

426 for Parvo6 was asparagine, specific to CPV-2a viruses, and for Parvo9 was aspartic acid which is specific to CPV-2b viruses. The amino acid residue at position 297 was alanine for both Parvo6 and Parvo9 suggesting they are new strains of CPV-2a and -2b, as described by Decaro et al. [27].

4.5 Discussion

The NGS methodology proved to be an effective diagnostic tool for further characterizing CPV in rectal swabs of affected canines. Not only was the virus detected, but the RNA for all viral proteins was obtained in many cases suggesting that the virus was undergoing active replication and causing symptoms. Despite only targeting RNA transcripts, near whole genomes were obtained making it possible to construct a phylogenetic tree that demonstrated the presence of 2 distinct subtypes of a CPV virus, CPV-2a in dogs living in and around Community A and CPV-2b in dogs living in Community B. Similarity of these 2 CPV virus groups to previously published CPV sequences as well as the limited presence of protective CPV antibodies in tested serum suggest that the outbreak did not involve a novel strain of CPV, but rather exposure in an under-vaccinated on unvaccinated, and therefore immunologically naive, dog population.

4.6 Funding

This research was supported in part by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM103395 and by an equipment grant of Alaska State Public Health Laboratories.

4.7 Competing Interests

The authors declare no competing interests or conflict of interest.

4.8 Acknowledgements

The authors wish to thank the veterinary community in Alaska for contributing specimens for further investigation as well as the laboratorians at the Biological Research and Diagnostics at the University of Alaska Fairbanks for performing initial testing and referral.

Table 4.1: Specimens and referral results summary

Sample ID	Description	Community (A or B)	Total Ab (HAI)	IgG only (HAI)	CPV PCR	Genotyping
Parvo1	8wk old	A	--	--	Positive	2/2a
Parvo2	Same animal as Parvo1 (reproducibility specimen)	A	--	--	Positive	2/2a
Parvo3	8wk old littermate to Parvo1/2	A	--	--	Positive	--
Parvo4	6 mo. old	B	--	--	Positive	--
Parvo5	1.5 yr. old	B	--	--	Positive	--
Parvo6	2 yr. old, moved from affected Community B to Community A	A	160	10	Positive	--
Parvo7	6 yr. old, next door neighbor to Parvo6. Died within one day of Parvo6.	A	80	<10	Positive	--
Parvo8	6 mo. old	B	--	--	--	--
Parvo9	Unknown	B	--	--	--	--
Parvo10	5 mo. old	B	--	--	--	2b
Parvo11	10 yr. old	B	--	--	Positive	--
Parvo12	20 mo. old	B	1,280	<40	Positive	--

-- = Not Tested, HAI = hemagglutination inhibition

Table 4.2: NGS metrics after aligning read files to reference genome NC_001539.1 (canine parvovirus)

Sample ID	Sequencing Output (Megabases)	Total Reads	Number of Viral Reads Aligned to Ref. genome	Viral Reads of Total Reads (%)	Read Depth (X)	Genome Coverage (%)
Parvo1	437.3	1,749,240	277	0.02	13.0	87.8
Parvo2	418.4	1,673,478	1,114	0.07	52.3	93.4
Parvo3	500.0	1,999,824	10,450	0.52	490.8	94.8
Parvo4	924.6	3,698,580	2,418	0.07	113.6	92.4
Parvo5	582.8	2,331,346	44,288	1.90	2,080.0	95.1
Parvo6	630.9	2,523,606	29,789	1.18	1,399.1	96.5
Parvo7	585.3	2,341,324	7,543	0.32	354.3	93.1
Parvo8	509.6	2,038,224	2,744	0.13	128.9	89.1
Parvo9	977.5	3,909,948	263,625	6.74	12,381.4	94.9
Parvo10	348.2	1,392,626	2,906	0.21	136.5	89.4
Parvo11	767.3	3,069,074	47	0.0015	2.2	44.8
Parvo12	435.3	1,741,200	3,439	0.20	161.5	93.5

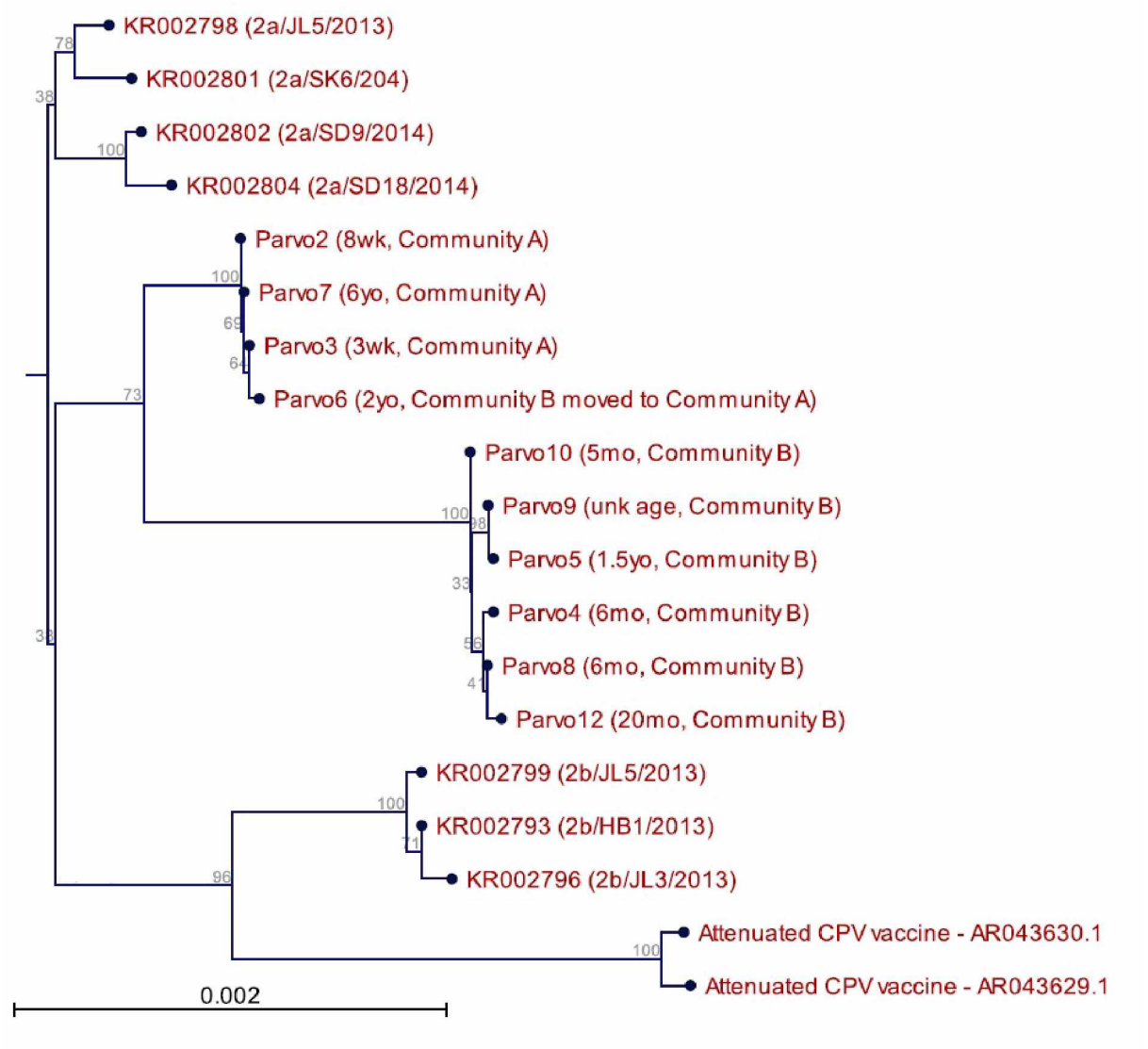
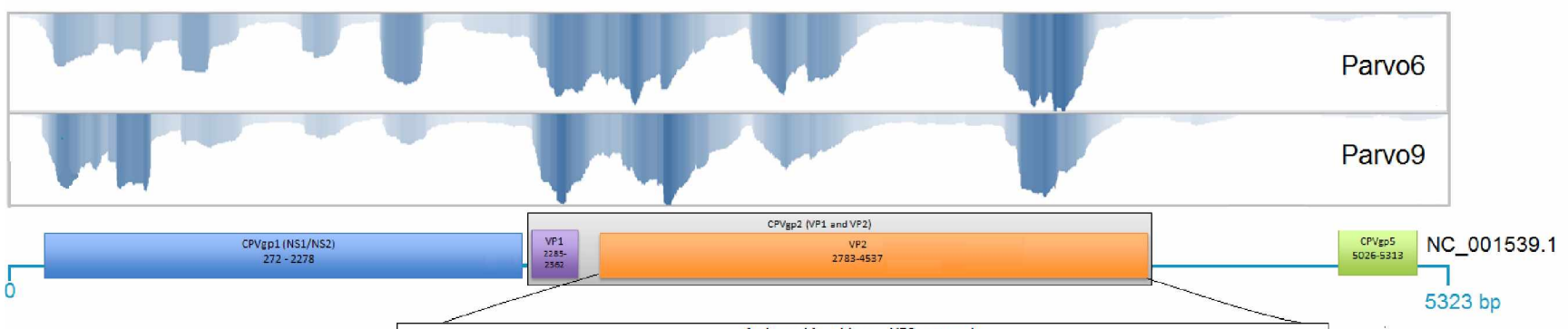


Figure 4.1: Phylogenetic tree of clinical specimen isolates in relation to reference sequences of canine parvovirus 2a, 2b, and vaccine candidate genomes. Parentheses next to 10 of the 12 specimens analyzed indicate the age of the canine and kennel origination. Two specimens, Parvo1 and Parvo11, are not included in this analysis due to inadequate read depth and subsequent insufficient sequence availability. Two major groupings are recognized as distinctly 2a viruses (Parvo2, 7, 3, 6) and 2b viruses (Parvo10, 9, 5, 4, 8, 12).



	Amino acid residues – VP2 gene only																	
	80	87	93	101	103	232	267	297	300	305	323	324	375	426	440	555	564	568
New CPV-2a viruses																		
Parvo2	R	L	N	T (ACC)	A	I	Y (TAT)	A	G	Y	N	I (ATT)	D	N (AAT)	A (GCA)	V	S	G
Parvo7	R	L	N	T (ACC)	A	I	Y (TAT)	A	G	Y	N	I (ATT)	D	N (AAT)	T (ACA)	--	--	--
Parvo3	R	L	N	T (ACC)	A	I	Y (TAT)	A	G	Y	N	I (ATT)	D	N (AAT)	T (ACA)	--	--	--
Parvo6	R	L	N	T (ACC)	A	I	Y (TAT)	A	G	Y	N	I (ATT)	D	N (AAT)	A (GCA)	V	S	G
New CPV-2b viruses																		
Parvo10	R	L	N	T (ACT)	A	I	F (TTT)	A	G	Y	N	Y (TAT)	D	D (GAT)	T (ACA)	V	S	G
Parvo9	R	L	N	T (ACT)	A	I	F (TTT)	A	G	Y	N	Y (TAT)	D	D (GAT)	T (ACA)	V	S	G
Parvo5	R	L	N	T (ACT)	A	I	F (TTT)	A	G	Y	N	Y (TAT)	D	D (GAT)	T (ACA)	V	S	G
Parvo4	R	L	N	T (ACT)	A	I	F (TTT)	A	G	Y	N	Y (TAT)	D	D (GAT)	T (ACA)	V	S	G
Parvo8	R	L	N	T (ACT)	A	I	F (TTT)	A	G	Y	N	Y (TAT)	D	D (GAT)	T (ACA)	--	--	G
Parvo12	R	L	N	T (ACT)	A	I	F (TTT)	A	G	Y	N	Y (TAT)	D	D (GAT)	T (ACA)	V	S	G
Indeterminate (insufficient data)																		
Parvo1	R	L	N	T (ACT)	A	I	Y (TAT)	A	G	Y	N	I (ATT)	D	--	T (ACA)	--	--	--
Parvo11	R	L	N	--	--	--	--	--	--	--	--	Y (TAT)	D	D (GAT)	--	V	--	--

Figure 4.2. Visual depiction of sequence alignment to reference genome, VP2 gene location within the canine parvovirus genome and analysis of 18 amino acid positions. Parvo6 and Parvo9 were chosen as representatives of each group of viruses to visualize NGS data and varying sequencing depths for each major protein. The VP2 region of each specimen was analyzed at key amino acid positions reflecting canine parvovirus subtype. Parvovirus specimens are ordered and categorized as they are depicted in Figure 4.1 (Phylogenetic tree). "New" canine parvovirus 2a and 2b stem from emerging viruses showing variability at the 426 aa position [27].

4.9 References

1. Hoelzer, K. and C.R. Parrish, *The emergence of parvoviruses of carnivores.* Vet Res, 2010. **41**(6): p. 39.

2. Reed, A.P., E.V. Jones, and T.J. Miller, *Nucleotide sequence and genome organization of canine parvovirus.* Journal of virology, 1988. **62**(1): p. 266-276.

3. Nandi, S. and M. Kumar, *Canine Parvovirus: Current Perspective.* Indian journal of virology : an official organ of Indian Virological Society, 2010. **21**(1): p. 31-44.

4. Pollock, R.V., *Experimental canine parvovirus infection in dogs.* Cornell Vet, 1982. **72**(2): p. 103-19.

5. Pollock, R.V. and M.J. Coyne, *Canine parvovirus.* Vet Clin North Am Small Anim Pract, 1993. **23**(3): p. 555-68.

6. Gordon, J.C. and E.J. Angrick, *Canine parvovirus: environmental effects on infectivity.* American journal of veterinary research, 1986. **47**(7): p. 1464-1467.

7. Crawford, C., *Management of Disease Outbreaks in Animal Shelters.* 2013.

8. Schmitz, S., et al., *Comparison of Three Rapid Commercial Canine Parvovirus Antigen Detection Tests with Electron Microscopy and Polymerase Chain Reaction.* Journal of Veterinary Diagnostic Investigation, 2009. **21**(3): p. 344-345.

9. Decaro, N., et al., *Long-term viremia and fecal shedding in pups after modified-live canine parvovirus vaccination.* Vaccine, 2014. **32**(30): p. 3850-3.

10. Mylonakis, M.E., Kalli, I., Rallis, T.S., *Canine parvoviral enteritis: an update on the clinical diagnosis, treatment, and prevention.* Veterinary Medicine: Research and Reports, 2016. **7**: p. 91-100.

11. Markovich, J.E., et al., *Effects of canine parvovirus strain variations on diagnostic test results and clinical management of enteritis in dogs.* J Am Vet Med Assoc, 2012. **241**(1): p. 66-72.

12. Morrow, W., *Canine parvovirus outbreak hits Interior Alaska.* The Daily News Miner, February 25, 2016.

13. The Associated Press, *Canine parvovirus outbreak reported in Alaska's interior.* The Juneau Empire, February 28, 2016.

14. Thomas, M., *Outbreak of Parvovirus Confirmed in Alaskan Dog Kennels.* The Nome Nugget, March 4, 2016.

15. Soma, T., et al., *Analysis of the VP2 protein gene of canine parvovirus strains from affected dogs in Japan.* Res Vet Sci, 2013. **94**(2): p. 368-71.

16. Wang, H.-C., et al., *Phylogenetic Analysis of Canine Parvovirus VP2 Gene in Taiwan.* Virus Genes, 2005. **31**(2): p. 171-174.

17. Zhang, R., et al., *Phylogenetic analysis of the VP2 gene of canine parvoviruses circulating in China.* Virus Genes, 2010. **40**(3): p. 397-402.

18. Zienius, D., et al., *Phylogenetic characterization of Canine Parvovirus VP2 partial sequences from symptomatic dogs samples.* Pol J Vet Sci, 2016. **19**(1): p. 187-96.

19. Amrani, N., et al., *Molecular epidemiology of canine parvovirus in Morocco.* Infect Genet Evol, 2016. **41**: p. 201-6.

20. Clegg, S.R., et al., *Molecular Epidemiology and Phylogeny Reveal Complex Spatial Dynamics in Areas Where Canine Parvovirus Is Endemic.* Journal of Virology, 2011. **85**(15): p. 7892-7899.

21. Kapil, S., et al., *Canine Parvovirus Types 2c and 2b Circulating in North American Dogs in 2006 and 2007.* Journal of Clinical Microbiology, 2007. **45**(12): p. 4044-4047.

22. Ohshima, T., et al., *Chronological analysis of canine parvovirus type 2 isolates in Japan.* J Vet Med Sci, 2008. **70**(8): p. 769-75.

23. Pérez, R., et al., *Phylogenetic and Genome-Wide Deep-Sequencing Analyses of Canine Parvovirus Reveal Co-Infection with Field Variants and Emergence of a Recent Recombinant Strain.* PLoS One, 2014. **9**(11).

24. Zhu, Y., et al., *Genome Sequence of a Canine Parvovirus Strain, CPV-s5, Prevalent in Southern China.* Genome Announcements, 2014. **2**(1).

25. Ge, F., et al., *Preferential Amplification of Pathogenic Sequences.* Sci Rep, 2015. **5**.

26. Parker, J.C., J., *Application of next generation sequencing for the detection of humanviral pathogens in clinical specimens.* Journal of Clinical Virology, 2017. **86**: p. 20-26.

27. Decaro, N. and C. Buonavoglia, *Canine parvovirus--a review of epidemiological and diagnostic aspects, with emphasis on type 2c.* Vet Microbiol, 2012. **155**(1): p. 1-12.

General Conclusions

The studies included in this thesis were aimed at comparing new technologies to conventional methods to help move clinical laboratories forward in their quest to adopt modern methods of viral pathogen detection. State and federal regulations, such as CLIA, regulate how these assays are validated and implemented for use by clinicians [1]. Published studies, such as these, help fulfill components of these validation requirements and can be used by clinical laboratories to reduce the time and cost of re-validating the same process in each individual laboratory and will also help enhance standardization.

I began by describing the process of validating an assay's analytical sensitivity. The issue of whether or not extensive multiplexed respiratory pathogen panels perform at the same level of sensitivity when compared to their singleplex counterparts was explored. Chapter 1 evaluates a commercial assay that utilizes electrochemical detection of PCR amplicons, a novel technology produced by GenMark Diagnostics. This multiplex assay was compared to established singleplex assays in terms of analytical sensitivity, or lower limit of detection. This evaluation assists other clinical laboratories that struggle with making choices between utilizing available commercial technology and continuing on with laboratory-developed tests [2]. This experiment identified that the analytical sensitivities are challenging to maintain in multiplexed systems. In contrast, some assays showed better performance when multiplexed. It was determined that the analytical sensitivity of the electrochemical technology showed adequate sensitivity for most respiratory viruses tested except for the generic influenza A assay. Clinical laboratories are dependent on sensitive generic influenza A assays for screening clinical specimens prior to subtyping. Because the screening assay was found to be less sensitive, false negativity may be an issue. This is especially concerning for influenza surveillance centers aimed

at targeting the occurrence of influenza A viruses due to its history of causing devastating pandemics across species. Although this technology proved to be comparable in terms of the sensitivity of most respiratory viruses tested, users should caution the lack of sensitivity for influenza A detection when compared to singleplex assays.

Chapter 2 aims at measuring the positive predictive value (PPV) of an assay performed using new technology when imposed on a population of low disease prevalence, for example, when conducting HIV surveillance in Alaska. This is an important topic because assays, in general, perform poorly in terms of false positivity when used in populations with limited disease. An example of an assay evident of this phenomenon is rapid tests for influenza being widely used in hospitals and clinics. When influenza prevalence is low in a population, false positivity occurs more often on rapid tests. On the other hand, when influenza prevalence is high like in the middle of respiratory virus season, the rate of false positivity goes down. In terms of HIV, the assay examined uses multiplex flow immunoassay technology marketed by BioRad Laboratories to broadly capture the presence of HIV-1 and HIV-2 viruses and/or antibodies in a patient's serum. It was determined that the PPV was acceptable (62.9%) when compared to other assays currently on the market. Also, it should not be understated that this new technology offers broad detection of HIV viruses while simultaneously providing differentiation of which type of virus and/or antibody is present. This can be extremely advantageous for interpreting test outcomes and relating it to a patient's HIV status, even when faced with less than desirable false positivity rates.

Chapter 3 describes the potential of using NGS compared to conventional methods for virus detection. Using turnaround time and result outcomes as an indicator of clinical utility clearly reveal that NGS is capable of fully characterizing viral genomes, even when heavily

masked in the background of the specimen metagenome [3]. Full characterization not only identified the isolate of the viral species but also identified known antiviral resistance markers, both of which have significant clinical impact on disease diagnosis and treatment. Many applications in clinical laboratories using NGS are focused on targeted approaches to try to limit the amount of unwanted metagenomic data [4, 5]. The experiment described in Chapter 3 is the first of its kind describing the true fraction of targeted pathogen within the metagenome, without utilizing enrichment techniques, demonstrating the true scarcity of pathogenic nucleic acid in clinical specimens and the sensitivity potential of massively parallel sequencing.

Chapter 4 demonstrates the ability of NGS to not only detect and fully characterize a virus associated with an outbreak, but to investigate the outbreak source in detail. In this example, rectal swabs were used to sequence two strains of canine parvovirus, occurring simultaneously in two different geographical locations. This study helped distinguish these viruses as similar to vaccine strains, not novel strains as previously assumed due to the amount of dog deaths associated with the outbreak. Phylogenetic evaluation of these strains in combination with serological data helped drive the conclusion that these dogs were most likely experiencing a disease outcome associated with un- or under-vaccinated dog populations [5-7]. This type of epidemiological analysis can be extrapolated to human viral pathogen outbreaks in order to associate cases of infection to their potential source, drive better outbreak interventions, and diminish community panic [6].

In conclusion, this thesis examines new technologies designed to detect and characterize pathogenic viruses. Some technologies, like those offered by companies like GenMark and BioRad, are well-developed and provide current benefit to the health of patients. Others, like next generation sequencing methods, require much more validation before becoming a routine

method in clinical laboratories. However, successful use of next generation sequencing methods in context of an infectious disease laboratory may better prepare clinical labs in their role in precision medicine, the long-term research goal of the National Institute of Health which focuses on understanding individual variability in disease prevention, care, and treatment [8]. Infectious disease laboratories should strive to adopt new, target-independent technologies to expand pathogen detection to support these types of initiatives.

References

1. Centers for Disease Control and Prevention (CDC) (2) Centers for Medicare & Medicaid Services (CMS), H., *Medicare, Medicaid, and CLIA programs; laboratory requirements relating to quality systems and certain personnel qualifications. Final rule.* Fed Regist, 2003. **68**(16): p. 3639-714.

2. Merrick, R., S.H. Hinrichs, and M. Meigs, *Public Health Laboratories*, in *Public Health Informatics and Information Systems*, J.A. Magnuson and J.P.C. Fu, Editors. 2014, Springer London: London. p. 295-308.

3. Hall, R.J., et al., *Evaluation of rapid and simple techniques for the enrichment of viruses prior to metagenomic virus discovery.* Journal of Virological Methods, 2014. **195**: p. 194-204.

4. Ge, F., et al., *Preferential Amplification of Pathogenic Sequences.* Sci Rep, 2015. **5**.

5. Gardy , J.L., et al., *Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak.* New England Journal of Medicine, 2011. **364**(8): p. 730-739.

6. Grad, Y.H., et al., *Genomic epidemiology of the Escherichia coli O104:H4 outbreaks in Europe, 2011.* Proceedings of the National Academy of Sciences, 2012. **109**(8): p. 3065-3070.

7. Lienau , E.K., et al., *Identification of a Salmonellosis Outbreak by Means of Molecular Sequencing.* New England Journal of Medicine, 2011. **364**(10): p. 981-982.

8. Collins, H., et al., *Information Needs in the Precision Medicine Era: How Genetics Home Reference Can Help.* Interactive journal of medical research, 2016. **5**(2): p. e13-e13.

Appendix A

Correction: Analytical Sensitivity Comparison between Singleplex Real-Time PCR and a Multiplex PCR Platform for Detecting Respiratory Viruses[5]

Abstract

Multiplex PCR methods are attractive to clinical laboratories wanting to broaden their detection of respiratory viral pathogens in clinical specimens. However, multiplexed assays must be well optimized to retain or improve upon the analytic sensitivity of their singleplex counterparts. In this experiment, the lower limit of detection (LOD) of singleplex real-time PCR assays targeting respiratory viruses is compared to an equivalent panel on a multiplex PCR platform, the GenMark eSensor RVP. LODs were measured for each singleplex real-time PCR assay and expressed as the lowest copy number detected 95–100% of the time, depending on the assay. The GenMark eSensor RVP LODs were obtained by converting the TCID50/mL concentrations reported in the package insert to copies/μL using qPCR. Analytical sensitivity between the two methods varied from 1.2–1280.8 copies/μL (0.08–3.11 log differences) for all 12 assays compared. Assays targeting influenza A/H3N2, influenza A/H1N1pdm09, influenza B, and human parainfluenza 1 and 2 were most comparable (1.2–8.4 copies/μL, <1 log difference). Largest differences in LOD were demonstrated for assays targeting adenovirus group E, respiratory syncytial virus subtype A, and a generic assay for all influenza A viruses regardless of subtype (319.4–1280.8 copies/μL, 2.50–3.11 log difference). The multiplex PCR platform, the GenMark eSensor RVP, demonstrated improved analytical sensitivity for detecting influenza

[5] Original publication as cited, including correction in Table 3: Parker, J., N. Fowler, M. L. Walmsley, T. Schmidt, J. Scharrer, J. Kowaleski, T. Grimes, S. Hoyos and J. Chen. *Correction*: *Analytical Sensitivity Comparison between Singleplex Real-Time PCR and a Multiplex PCR Platform for Detecting Respiratory Viruses*. PLOS ONE, 2015. **10**(11): e0143164.

A/H3 viruses, influenza B virus, human parainfluenza virus 2, and human rhinovirus (1.6–94.8 copies/μL, 0.20–1.98 logs). Broader detection of influenza A/H3 viruses was demonstrated by the GenMark eSensor RVP. The relationship between TCID50/mL concentrations and the corresponding copy number related to various ATCC cultures is also reported.

Introduction

Multiplex PCR methods, those that target more than one pathogen in a single test, benefit diagnostics in a clinical laboratory due to their ability to detect and rule-out many related pathogens in the same amount of time. New and improved workflow designs make it possible for laboratories with varied molecular technical ability to implement multiplex PCR platforms.

The Respiratory Viral Panel (RVP) manufactured by GenMark Diagnostics, Inc. is a multiplex PCR panel that detects the amplification of various viral gene fragments electrochemically. Nucleic acids from targeted viral pathogens are amplified using a multiplex PCR reaction followed by denaturation of the double stranded molecules into single oligonucleotide strands using exonuclease. Once the amplicons are in a single-stranded state, they are hybridized to a complementary virus-specific signal probe tagged with ferrocene, a reducing agent. This hybridized molecule is then exposed to another sequence-specific probe which is bound to a solid phase, a gold electrode. Upon application of a low voltage current, the hybridized molecule bound to this solid phase brings the ferrocene in close proximity to the gold electrode where reversible electron transfer can occur and the resulting current can be measured. Viral pathogenic nucleic acid can be detected with confidence when measurements are at or exceed 3 nanoamps (nA) on the GenMark XT-8 instrument. The GenMark eSensor RVP has been shown to be highly comparable to other multiplex PCR platforms as well as singleplex real-

time PCR in terms of diagnostic sensitivity and specificity[1,2], which measures the level of correlation between two methods. In this experiment, the primary interest is the analytical sensitivity of the PCR assays, or the minimum detectable concentration of the target. The GenMark eSensor RVP LODs as determined by the manufacturer are compared to singleplex real-time PCR assay LODs determined by our laboratory and expressed as lowest copy number reliably detected 95–100% of the time.

Limit of detections for FDA-approved clinical assays, including those described in the GenMark eSensor RVP package insert, are commonly expressed as 50% tissue culture infectious dose per milliliter, or TCID50/mL. Although this is a standard practice, other quantification methods such as real-time PCR are also reliable and may be able to more precisely describe quantities of viral particles with or without TCID50/mL calculations as a reference[3–6]. Since the LODs for the GenMark RVP assays are expressed exclusively as TCID50/mL concentrations, these values needed to be converted to copy number per μL in order to meet our goals of comparing analytical sensitivity as lowest copy number. The LODs of each GenMark RVP assay were not re-established in our laboratory. Instead, manufacturer established TCID50/mL values were converted to copy number using quantitative real-time PCR (qPCR). Performing this conversion also provided an opportunity to view the relationship between TCID50/mL and copy number and relate this information to various virus-infected ATCC cell cultures.

The respiratory assays evaluated in this experiment target the following virus species: influenza A virus (InfA/H3N2 and InfA/H1N1pdm09), influenza B virus (InfB), human respiratory syncytial virus (RSV), human parainfluenza virus (hPIV 1, 2, and 3), human adenovirus (Adeno), and human rhinovirus (hRV). The multiplex GenMark eSensor RVP assays

were able to further distinguish human adenoviruses as belonging to subgenera C or E and respiratory syncytial viruses as belonging to subgroup A or B, unlike the singleplex real-time PCR assays that were designed to detect human adenovirus and respiratory syncytial virus universally across all subgroups. A generic influenza A virus assay, one that targets a conserved region of all influenza A viruses regardless of strain, was also evaluated.

Methods and Materials

Clinical specimens

Clinical specimens used in this study were de-identified. The University of Alaska Fairbanks Institutional Review Board (IRB) has determined that the proposed research qualifies for exemption from the requirements of 45 CFR 46 (Approval number: 667418–1).

Preparation of standard materials

Specific plasmids were created for each real-time PCR assay by ligating single copies of the diagnostic amplicon onto vectors (pCR 2.1 or pCR4, Invitrogen) and amplifying via TOPO cloning (Invitrogen). Transformant E.coli competent cells were extracted using a phenol/chloroform mixture and the presence of viral-specific inserts was verified by sequencing (Elim Biopharmaceuticals, Inc.).

Plasmid concentrations were calculated by performing two quantification methods: 1) fluorometry specific to double stranded DNA (Qubit 2.0, dsDNA br Assay Kit, Invitrogen) and 2) pixel intensity measurements using the ImageJ application[7]. Using ImageJ, the pixel intensity of linearized plasmid DNA gel bands could be interpolated into a standard curve consisting of 1KB ladder dilutions (New England Biolaboratories) to predict quantities of unknown bands on the gel. Plasmid DNA was linearized using restriction enzyme NcoI (New

England Biolabs) prior to gel electrophoresis. These quantification strategies were chosen to focus on the DNA of interest and to help exclude possible quantification pitfalls of over or underestimating DNA concentrations. Used in combination, these methods accounted for contaminating RNA (fluorometry specific for DNA only) as well as contaminating DNA as seen as different sized bands on the gel which could be excluded by only measuring the pixel intensity of gel bands of expected size (~4KB).

Differences between the two quantification methods ranged from 0.2 to 5.4 ng/μL (average 2.6 ng/μL ± 1.8). Final concentrations were calculated by rounding the average of the two methods to the nearest 2.5ng/μL. The weight of each plasmid was calculated using Geneious (v.8.1.3), using the known sequence of the vector in addition to the confirmed sequence of the insert. Final copy numbers (per μL) were calculated by dividing the plasmid weights (ng/copy) into the concentrations of each plasmid (ng/μL). Results of the quantification methods and downstream calculations are shown in Table A.1.

Table A.1: Plasmid concentrations and copy number determination

Virus target insertion	Vector	Concentration (ng/μL)		avg ± SD (ng/ μL)	Final (ng/ μL)	Weight per plasmid copy* (ng)	Copies/μL
		Qubit	*ImageJ*				
Adeno	pCR2.1	8.0	7.8	7.9 ± 0.2	7.5	4.17×10^{-9}	1.80×10^{9}
InfA	pCR 2.1	10.1	11.5	10.8 ± 1.0	10	4.14×10^{-9}	2.42×10^{9}
InfA/H3N2	pCR4	14.5	9.1	11.8 ± 3.8	10	4.37×10^{-9}	2.29×10^{9}
InfA/H1N1pdm09	pCR2.1	11.1	11.8	11.5 ± 0.5	10	4.15×10^{-9}	2.41×10^{9}
InfB	pCR4	26.0	31.1	28.6 ± 3.6	30	4.17×10^{-9}	7.19×10^{9}
hPIV-1	pCR4	13.2	8.7	9.5 ± 2.0	10	4.32×10^{-9}	2.31×10^{9}
hPIV-2	pCR4	7.4	6.7	7.0 ± 0.5	7.5	4.15×10^{-9}	1.81×10^{9}
hPIV-3	pCR4	38.8	35.0	36.9 ± 2.7	35	4.19×10^{-9}	8.35×10^{9}
RSV	pCR4	5.1	6.2	5.7 ± 0.8	5	4.17×10^{-9}	1.20×10^{9}
hRV	pCR4	32.7	29.8	31.3 ± 2.1	30	4.27×10^{-9}	7.03×10^{9}

*Weight/copy was calculated using Geneious (v.8.1.3) which considers the exact sequence of the plasmid.

Determination of singleplex real-time LOD

Plasmid DNA was serially diluted to produce eight (8) test concentrations ranging between 1 copies/μL and 1250 copies/μL, depending on the assay. This narrow range was chosen to identify the lowest potential copy number able to be detected repeatedly, but keep it above theoretical limitations of real time PCR, <3 copies (0.6 copies/uL when using 5uL per reaction)[8]. Seven (7) replicates were tested at each concentration. This process was repeated twice, once using nuclease-free water as the diluent background for the plasmids to assess basic analytical sensitivity and once using total nucleic acid extract (TNA) as background for the plasmids to simulate real clinical matrices. TNA was isolated from clinical specimens using the easyMAG total nucleic acid automated extractor (Biomerieux). A total of 200μL of the clinical specimen was extracted and final eluate volumes were 60μL. TNA from clinical specimens were screened by PCR, and only those that demonstrated the absence of target DNA or RNA were qualified to be pooled as clinical background diluent.

Primers and probes used in the laboratory-developed real-time PCR assays have been previously described[9,10]. Influenza assays were performed using Invitrogen Superscript III reagents and all other assays were performed using Ambion AgPath ID reagents. For assays using the Invitrogen reagents, the following PCR thermal cycling profile was used; 50°C hold for 30 minutes, 95°C hold for 2 minutes, and 45 cycles of 95°C for 15 seconds then 55°C for 30 seconds. For assays using the Ambion reagents, the following PCR thermal cycling profile was used; 45°C hold for 10 minutes, 95°C for 10 minutes, and 45 cycles of 95°C for 15 seconds then 55°C for 1 minute. Reactions were tested using ABI 7500Dx thermal cyclers (Life Technologies).

Negative controls consisted of no template control replicates (NTC, n = 3) and diluent blank replicates, made up of water or TNA diluent (n = 7) to assess contamination. Positive reactions were defined as those amplification curves that produced cycle threshold (Ct) values at or below 40 cycles. The LOD was chosen as the concentration that demonstrated a percentage of positivity over all replicates at a particular dilution. The percentage of positivity was chosen using those that were set by the manufacturer for each matching GenMark RVP assay. All but three assays were set by the manufacturer below 100% positivity (InfA/H1N1pdm09, RSVA, and hRV assays only); therefore, the LOD for these particular singleplex assays were estimated using probit analysis to match these probabilities for comparison purposes[11]. Final LODs were expressed as a concentration, copies/µL (Table A.2).

Table A.2: LOD comparison summary

		Singleplex Real-time PCR		Multiplex PCR GenMark eSensor RVP		
		Lowest copies/µL detected				
Assay	%pos	*Clinical background*	*No Background*	copies/µL equivalent of $TCID_{50}$/mL LOD	copies/µL difference	Log Difference
Adeno C	100%	1.6	4	110.4 ± 8	108.8	2.04
Adeno E	100%	1.6	4	390.4 ± 45.4	388.8	2.59
InfA	100%	5.4	21.2	1286.2 ± 23.2	1280.8	3.11
InfA/H3N2	100%	10.6	21.2	<2.2	8.4	0.92*
InfA/H1N1	97.5%	7	7	10 ± 4.4	3	0.48
InfB	100%	2.6	53.2	1 ± 2.8	1.6	0.20*
hPIV-1	100%	1	1	<2.2	1.2	0.08
hPIV-2	100%	5.4	2.2	1.6 ± 0.6	3.8	0.58*
hPIV-3	100%	2.2	1	134.8 ± 8.4	132.6	2.12
RSVA	97.5%	6.8	3.6	326.2 ± 22.8	319.4	2.50
RSVB	100%	10.6	5.2	120.2 ± 8.6	109.6	2.04
hRV	95%	111.8	82.4	<17	94.8	1.98*

*lower LOD demonstrated for the multiplex assay; 5 μ L used in each reaction. Adenovirus and RSV assays were not differentiated with the singleplex real-time PCR assay, although RSV assays were calculated differently based on %pos to be compared. The $TCID_{50}$/mL concentration for InfA/H3, HPIV 1, and hRV exceeded the detection limit on the qPCR assay. Copy number difference was calculated by subtracting the lowest copies/ μ L detected with clinical background on the singleplex assays from the average copies/ μ L equivalent converted from $TCID_{50}$/mL.

Conversion of TCID50/mL concentrations to copies/µL

Cell cultures with known TCID50/mL quantities of target viruses (ATCC) were used to estimate the LOD for the GenMark RVP assay. Cultures were stored in liquid nitrogen until they were extracted using the easyMAG total nucleic acid automated extractor (Biomerieux). A total of 200µL of the TCID50/mL culture was extracted and final eluate volumes were 60µL. Purified nucleic acid was stored at -80°C until tested by quantitative real time PCR (qPCR).

Using quantified plasmids containing inserts specific to each assay, ten-fold dilutions were prepared covering 101 to 106 copies/5µL. Each dilution was tested in triplicate to create a standard curve. All qPCR assays utilized a sequence-specific hydrolysis probe with the exception of the H3 due to sequence incompatibilities with the ATCC strain being analyzed (see results). In this case, a SYBR Green assay (GoTaq, Promega) with new primers were designed to target this specific strain of Influenza A/H3. Alongside the standard curve, dilutions of the isolated nucleic acid derived from the ATCC cultures were tested in triplicate at dilutions that would include reported GenMark eSensor RVP LOD TCID50/mL values. As with the singleplex real-time PCR assays, reactions were tested on ABI 7500Dx thermal cyclers (Life Technologies) and standard curves and associated unknown quantities were calculated using ABI 7500 v2.3 software. The copy number equivalents for each GenMark eSensor RVP assay's LOD is shown in Table 2. The relationship between copy number and TCID50/mL for each ATCC culture tested is shown in Table 3.

Table A.3: Relationship between TCID50/mL concentrations and copy number

ATCC Culture		Genome Copies/$TCID_{50}$ (± SD)	LOD for GenMark eSensor assays ($TCID_{50}$/mL)
VR-1	Adenovirus Type 1 (C)	7 ± 1	8.89×10^{1}
VR-1572	Adenovirus Type 4 (E)	124 ± 14	1.58×10^{1}
VR-547	Influenza A/H3 (Aichi)	0.01 ± 0	1.58×10^{3}
VR-1736	Influenza A/H1N1	2,381 ± 1,048	1.05×10^{-1}
VR-101	Influenza B	16 ± 44	3.16×10^{-1}
VR-94	Human Parainfluenza Virus Type 1 (C35)	391 ± 0	2.81×10^{-2}
VR-92	Human Parainfluenza Virus Type 2 (Greer)	0.03 ± 0.01	2.81×10^{0}
VR-93	Human Parainfluenza Virus Type 3 (C243)	24 ± 1	2.81×10^{1}
VR-1540	Respiratory Syncytial Virus (A2)	6 ± 1	2.81×10^{0}
VR-955	Respiratory Syncytial Virus (B9320)	38 ± 3	1.58×10^{0}
VR-483	Rhinovirus 3 FEB	53,797 ± 0	1.58×10^{-3}

SD = standard deviation, SD could not be calculated for VR-547, VR-94, and VR-483 since the $TCID_{50}$/mL concentration exceeded the detection limit on the qPCR assay.

Results

Ten singleplex real-time PCR assays were compared in terms of analytical sensitivity to twelve multiplex assays on the GenMark eSensor RVP. This difference stems from the fact that the singleplex real-time PCR assays are not designed to distinguish between different subgenera of human adenovirus or different subtypes of respiratory syncytial viruses (RSV), while the GenMark eSensor RVP differentiates between human adenovirus C and E as well as RSV subtype A and B. Thus, two additional assays were evaluated for the GenMark eSensor RVP. Analytical sensitivity was expressed as lowest copies/µL concentration for all assays.

The Genmark eSensor RVP capable of distinguishing between different subgenera of adenoviruses (C vs. E) demonstrated less analytical sensitivity than the generic singleplex real-time PCR assay targeting all adenoviruses, differing by 108.8 copies/µL (2.04 log difference), and 388.8 copies/µL (2.59 log difference), respectively. The difference in sensitivity may be due

to slight variations in the targeted priming region. The singleplex real-time PCR assays use primers designed to anneal highly conserved sequences within the hexon-coding region in order to target all adenoviruses, whereas the GenMark eSensor RVP assays use subgenera-specific hexon primers to make possible the distinction between adenovirus subgenera C and E. Upper respiratory tract infections associated with adenovirus C viruses infect more than 80% of the population early in life[12]; however, infections with the adenovirus E (serotype 4) can prove to be more severe and even fatal for people living in close quarters, such as military recruits[13]. In terms of surveillance, differentiation of virus subgenera within a population may be clinically useful, regardless of lost sensitivity.

Similarly, the singleplex real-time PCR assay generically targeting respiratory syncytial viruses also demonstrated better sensitivity than the GenMark eSensor RVP assays which are capable of distinguishing between subtypes A and B (319.4 copies/μL, 2.50 log difference and 109.6 copies/μL, 2.04 log difference, respectively). Respiratory syncytial viruses in subtype A are thought to be more prevalent and virulent than those in subtype B[14]. Subtyping respiratory syncytial virus may be clinically beneficial when surveilling populations that experience high hospitalization rates associated with the virus, such as Native Americans living in southwest United States and Alaska[15].

Analytical sensitivity of assays targeting the current circulating strains of influenza A viruses in the human population, H3N2 and H1N1pdm09, were highly comparable between the singleplex real-time PCR and multiplex GenMark eSensor RVP assays (8.4 copies/μL, 0.92 log difference and 3 copies/μL, 0.48 log difference, respectively). Comparing the LOD between the influenza H3N2 assays proved to be the most challenging. When converting TCID50/mL concentrations to copies/μL using qPCR, it was determined that this particular culture contained

an uncommon virus, an Aichi strain (A/Aichi/2/35) circa 1968 (ATCC) and therefore could not be amplified using the singleplex real-time PCR assay, which is designed to detect current influenza A/H3N2 virus strains. However, it was repeatedly detected using the GenMark eSensor RVP. This finding suggests that the eSensor RVP is capable of detecting a broader range of Influenza A/H3N2 strains while maintaining a comparable analytic sensitivity to that of its singleplex real-time PCR counterpart.

The greatest difference measured between analytic sensitivities was seen with the generic influenza A assay showing a 3.11 log difference in LOD (1280.8 copies/μL difference). Because the LOD for the generic influenza A assay is much higher than the subtype assays (as described above) for the multiplex GenMark eSensor RVP, difficulty in result interpretation from specimens with low influenza A virus titers is likely, since subtypes (H3N2 or H1N1pdm09) have a lower LOD than the generic influenza A assay (e.g. + H3N2,—influenza A). The performance of the generic influenza A assay is an important surveillance tool for tracking genetic changes among influenza A viruses. For instance, specimens demonstrating positivity for influenza A using this generic, highly conserved matrix-coding region may not subtype using the H3N2 or H1N1pdm09 assays, which may indicate that the virus is novel and worthy of alerting public health authorities. In contrast, the influenza B assays were shown to be highly comparable between the singleplex and multiplex assays, with a difference of only 1.6 copies/μL (0.20 log difference).

Human parainfluenza 1 assays were highly comparable (1.2 copies/μL, 0.08 log difference). Human parainfluenza 2 assays demonstrated improved sensitivity on the multiplex GenMark eSensor assay (3.8 copies/μL, 0.58 log difference). Human parainfluenza 3 assays demonstrated the largest difference in analytical sensitivity among the human parainfluenza

serotypes, demonstrating a 2.12 log improvement in detectability when using the singleplex real-time PCR assay (132.6 copies/μL difference).

Five of the twelve GenMark eSensor RVP assays matched (<1 log difference in copies/μL) the LOD of the real-time singleplex PCR assay targets in this study (Table 2). These include influenza A/H3N2, influenza A/H1N1pdm09, influenza B, and human parainfluenza 1 and 2. Six of the twelve assays compared showed greater sensitivity using the real-time singleplex assays. These include the adenovirus assays (C & E), influenza A, human parainfluenza 3, and RSV (A & B). The GenMark eSensor human rhinovirus assay demonstrated the biggest difference in terms of improved detection when compared to its singleplex counterpart (94.8 copies/μL, 1.98 log difference, 95% positivity).

The number of genome copies per TCID50/mL value was highly variable ranging from 0.01 to 53,797 (Table 3). LODs set at higher TCID50/mL concentrations (102–103) corresponded to stock cultures with lower copy numbers (0.01 to 6 copies). LODs set at in the mid-range TCID50/mL concentrations (101 to 10−1) corresponded to stock cultures with variable copy numbers per TCID50/mL (7–2,381 copies). LODs set at lower TCID50/mL concentrations (10−2–10−3) corresponded to stock cultures with somewhat higher copy numbers per TCID50/mL (391–53,797 copies).

Conclusion

Multiplex PCR applications benefit diagnostics in a clinical laboratory due to their ability to detect and rule-out many related pathogens in a single reaction, reducing tech-time by more than 3 hours for a panel of 10 viruses[1]. However, multiplex PCR platforms continue to carry higher overall costs. Analytic sensitivity, or the lowest possible concentration necessary to

produce a reliable result, is an important parameter to consider when replacing singleplex real-time PCR assays with multiplex PCR platforms evolving from newer, more expensive technologies. This experiment aims at finding a method in which to compare LODs of various assays using copy number as the unit of expression.

Choosing a 2.5 log difference to express considerable loss in sensitivity, the multiplex PCR strategy in combination with the GenMark eSensor technology demonstrates a considerable loss in sensitivity for three of the twelve assays assessed. Two of the assays were adenovirus E and respiratory syncytial virus subtype A. Although sensitivity is reduced, further characterization of viruses in clinical specimens may be of greater clinical importance, especially when particular subtypes are known to be more virulent in the population as is the case with adenovirus serotype 4 (subgenera E) and respiratory syncytial virus subtype A in particular populations.

The third assay demonstrating considerable loss in sensitivity was for the generic influenza A assay. Clinical laboratories, especially those directly related to public health surveillance, may need to consider the significance of this reduced sensitivity since it is commonly used to rule out novel influenza. Better analytic sensitivity was achieved using singleplex real-time PCR, which indicates that influenza A can be detected in clinical specimens even at low titers using this method. Specimens collected from patients that are suspected to have influenza infections that test negative on the GenMark eSensor RVP may need to be tested by more sensitive methods to rule out cases of novel influenza.

Expressing LOD in units that can be comparable across methodologies can prove to be difficult experimentally. TCID50/mL measurements can vary depending on how these cultures are handled in the laboratory in regards to preserving the concentration of infectious virus

particles for purposes of experimentation and quantity comparisons. Molecular detection strategies used in clinical laboratories are non-discriminating when identifying infectious or non-infectious viruses. PCR methodologies used to detect viral targets in clinical specimens do not provide information regarding the viability of the virus and, therefore, every detection may not point to a causative agent of disease. Other complicating factors to consider when interpreting PCR results are that patients can be asymptomatic carriers or may be exhibiting evidence of a past infections. Viral copy numbers provide an estimate of the number of virus particles in a given volume, but in our experiment, they did not correlate well with the number of infectious particles. To test the analytical sensitivity of a PCR-based methodology, it is important to understand that the intent of the assay is to detect any genome copy targeted by the designed primers, whether these be from infectious or non-infectious virus particles.

Acknowledgments

This research was supported by the Alaska Department of Health and Social Services, Division of Public Health, Section of laboratories. Much of the plasmid development and sequencing was supported in part by the University of Alaska Fairbanks. The ATCC cultures were purchased by GenMark Diagnostics, Incorporated in an effort to be consistent with the particular strains used in the FDA validation testing. We would like to thank the staff at the Alaska State Virology Laboratory for all of their help with carrying out testing for this project.

Funding Statement

This research was supported by the Alaska Department of Health and Social Services, Division of Public Health, Section of laboratories. Much of the plasmid development and

sequencing was supported in part by the University of Alaska Fairbanks. The ATCC cultures were purchased by GenMark Diagnostics, Incorporated in an effort to be consistent with the particular strains used in the FDA validation testing.

References

1. Pierce VM, Hodinka RL (2012) Comparison of the GenMark Diagnostics eSensor Respiratory Viral Panel to Real-Time PCR for Detection of Respiratory Viruses in Children. Journal of Clinical Microbiology 50: 3458–3465. 10.1128/JCM.01384-12

2. Popowitch EB, O'Neill SS, Miller MB (2013) Comparison of the Biofire FilmArray RP, Genmark eSensor RVP, Luminex xTAG RVPv1, and Luminex xTAG RVP Fast Multiplex Assays for Detection of Respiratory Viruses. Journal of Clinical Microbiology 51: 1528–1533. 10.1128/JCM.03368-12

3. Jonsson N, Gullberg M, Lindberg AM (2009) Real-time polymerase chain reaction as a rapid and efficient alternative to estimation of picornavirus titers by tissue culture infectious dose 50% or plaque forming units. Microbiology and Immunology 53: 149–154. 10.1111/j.1348-0421.2009.00107.x

4. Gustafsson RKL, Engdahl EE, Fogdell-Hahn A (2012) Development and validation of a Q-PCR based TCID(50) method for human herpesvirus 6. Virology Journal 9: 311–311. 10.1186/1743-422X-9-311

5. Kallesh D, Hosamani M, Balamurugan V, Bhanuprakash V, Yadav V, et al. (2009) Quantitative PCR: A quality control assay for estimation of viable virus content in live attenuated goat pox vaccine. Indian journal of experimental biology 47: 911

6. Iwami S, Holder BP, Beauchemin C, Morita S, Tada T, et al. (2012) Quantification system for the viral dynamics of a highly pathogenic simian/human immunodeficiency virus based on an in vitro experiment and a mathematical model. Retrovirology 9: 18 10.1186/1742-4690-9-18

7. Rasband WS ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997–2014.

8. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clinical Chemistry 55: 611–622. 10.1373/clinchem.2008.112797

9. Weinberg GA, Schnabel KC, Erdman DD, Prill MM, Iwane MK, et al. (2013) Field evaluation of TaqMan Array Card (TAC) for the simultaneous detection of multiple respiratory viruses in children with acute respiratory infection. J Clin Virol 57: 254–260. 10.1016/j.jcv.2013.03.016

10. Wangchuk S, Thapa B, Zangmo S, Jarman RG, Bhoomiboonchoo P, et al. (2013) Influenza surveillance from November 2008 to 2011; including pandemic influenza A(H1N1)pdm09 in Bhutan. Influenza and Other Respiratory Viruses 7: 426–430. 10.1111/j.1750-2659.2012.00409.x

11. Sloan LM (2007) Real-time PCR in clinical microbiology: verification, validation, and contamination control. Clinical Microbiology Newsletter 29: 87–95.

12. Garnett CT, Erdman D, Xu W, Gooding LR (2002) Prevalence and Quantitation of Species C Adenovirus DNA in Human Mucosal Lymphocytes. Journal of Virology 76: 10608–10616.

13. Robert NP, Joyce AC, Craig TM, Joel CG (2012) Adenovirus-associated Deaths in US Military during Postvaccination Period, 1999–2010. Emerging Infectious Disease journal 18: 507.

14. Walsh EE, McConnochie KM, Long CE, Hall CB (1997) Severity of Respiratory Syncytial Virus Infection Is Related to Virus Strain. Journal of Infectious Diseases 175: 814–820.

15. Holman RC, Curns AT, Cheek JE, Bresee JS, Singleton RJ, et al. (2004) Respiratory syncytial virus hospitalizations among American Indian and Alaska Native infants and the general United States infant population. Pediatrics 114: e437–444.

Appendix B

Application of next generation sequencing for the detection of human viral pathogens in clinical specimens[6]

Abstract

Background: Next generation sequencing (NGS) is a new technology that can be used for broad detection of infectious pathogens and is rapidly becoming an essential platform in clinical laboratories. It is not known how NGS will displace or enhance gold standard methodologies in infectious disease diagnosis.

Objectives: To investigate the feasibility and application of NGS technology in public health laboratories and compare NGS technology with conventional methods.

Study Design: Illumina MiSeq system was used to detect viral pathogens alongside other conventional virology methods using typical clinical specimen matrices. Sixteen clinical specimens and two CDC proficiency panels containing seventeen specimens were analyzed.

Results: Known pathogenic viral nucleic acid was positively identified in all clinical specimens, correlating and building upon results obtained by more conventional laboratory methods. Sequencing depths ranged from 0.008X to 319 and genome coverage ranged from 0. 6% to 99.9%. To substantiate the described methods used to analyze data derived from clinical specimens, the results of a clinical proficiency panel are also presented.

Discussion: Our results reveal true scarcity of known pathogenic viral nucleic acids in clinical specimens. NGS outperforms more conventional detection methods in this study by turnaround time as well as the improved depth of knowledge in regards to serotyping and drug resistance.

[6] Original publication as cited: Parker J, Chen J. *Application of next generation sequencing for the detection of human viral pathogens in clinical specimens.* Journal of Clinical Virology. 2016; **86**:20-26

Background:

Methodologies to detect pathogenic viruses in clinical specimens have transitioned from classic cell culture and antibody-antigen techniques to more sensitive molecular methods such as polymerase chain reaction (PCR). The targeted nature of these methodologies inhibit their ability to accommodate the true diversity of human pathogens in a clinical specimen, especially viruses [1]. Next generation sequencing (NGS) technologies are quickly demonstrating their ability to provide broad detection of infectious agents in a target-independent manner [2-7]. NGS has many advantages beyond the improved detection of all suspected, unsuspected, or even novel pathogens in a clinical specimen [8]. Familiarization with pathogen genomic sequences within clinical specimens enhances our understanding of infectious disease through further discovery of pathogen variability and genotyping [9-13], drug resistance or response to therapy [14-16], vaccine development and efficacy monitoring [17], and further characterization of the metagenome [18, 19]. The use of NGS for routine use in clinical diagnostics is emerging with its own set of limitations and challenges [13, 20]. Focusing on viruses of public health importance, we compared the performance of NGS alongside other more common viral detection methodologies.

Objectives

- To investigate the feasibility and application of NGS technology in public health laboratories and compare NGS technology with conventional methods
- To examine genome coverage and read depth of viral nucleic acid in various types of clinical specimens.

- To address the need for acceptability standards when using NGS due to the true scarcity of pathogenic viral nucleic acids in some clinical specimens.

Study Design:

Specimens: Sixteen previously tested clinical specimens, swab and serum specimens, were provided by the Alaska State Virology Laboratory in Fairbanks, Alaska. Two proficiency panels with a combined seventeen specimens for detecting antiviral resistance markers in the neuraminidase gene of influenza A virus were also tested as a quality indicator of our process. Proficiency specimens consisted of cultured Madin-Darby Canine Kidney Epithelial (MDCK) cells infected with influenza A virus.

Construction of sequencing library: Nucleic acid was isolated from 500µL of the original clinical specimen using phenol/chloroform followed by ethanol precipitation. DNA or RNA molecules were selected for by using DNase I (serum and proficiency specimens) or RNAse (swab specimens, with the exception of the influenza specimens). Quantity was evaluated using the Qubit instrument (Thermo Fisher Scientific) and the Agilent Bioanalyzer. The Nextera DNA Specimen Preparation protocol (Illumina) and the NEBNext Ultra RNA Library Prep Kit protocol (New England Biolab) were followed to prepare sequencing libraries.

Sequencing and data analysis. Libraries underwent paired-end sequencing on the Illumina MiSeq using a v.2 500-cycle kit. Read files were imported into Pathseq™Virome for reference genome identification. Alignments to the identified viral genome sequence(s) were performed by Sequencher (v5.1) in addition to an external tool, Genomic Short-read Nucleotide Alignment Program (GSNAP) [21]. Read depth and genome coverage was established using Tablet (v.1.13.12.17, [22]).

Results

NGS for detecting clinical adenovirus infections

Adenoviruses are important to characterize in the laboratory since some serotypes are more commonly associated with outbreaks, severe pneumonia, and possibly cancer such as serotypes 14, 55, and 12 [25]. Two infections were able to be diagnosed and further characterized using NGS (Figure 1).

Specimen Type	Patient	Culture	Staining	Serum Neutralization Assay	NGS → TAT: 4 days
Nasopharyngeal Swab	2yo M	Hep#2 3+ CPE - Day 2	DFA, Adenovirus	Adenovirus (type 2) TAT: 14 days	Human adenovirus C strain human/ARG/A15932/2002/2 (Cidofovir susceptible)
Nasal Swab	29yo F	Hep#2 4+ CPE - Day 8	DFA, Adenovirus	Adenovirus (type 3) TAT: 18 days	Human adenovirus B strain human/USA/UFL_Adv3a51/2007/3

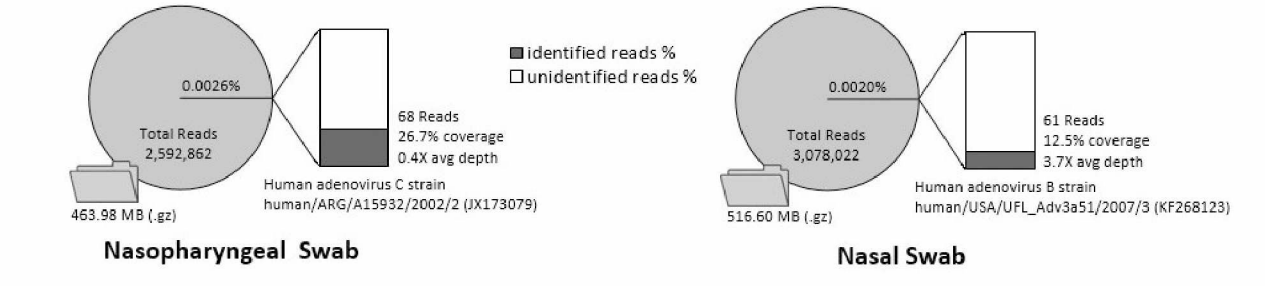


Figure B.1. Comparison of NGS and conventional virology assays for detecting adenovirus infection**.** Pie charts further describe sequence data as well as storage needed as compressed files (.gz). Percentage refers to the proportion of the total reads represented by the isolate reads. Exploding bar graph shows the percentage of the full genome that was identified by the reported number of reads.

NGS for detecting clinical herpesvirus infection

The results of three specimens are compared, two are clinical specimens and one is a cultured clinical isolate (Figure 2). As expected, the overall proportion of viral reads is much higher when sequencing clinical isolates (>15%) when compared to raw clinical materials (<1%).

Results were concurrent amongst all methods; however, the result obtained by conventional methods for the nasopharyngeal swab were not definitively concurrent. Sequence analysis identified human herpesvirus 5 strain HAN2 (JX512200), same as conventional methods, but by only 10 reads. Torque teno virus isolate US32 (AF122921) was more definitively identified. Surveillance for torque teno viruses is not common since they are thought to be ubiquitous in humans and lack concrete disease association [27]. Nevertheless, NGS alone could not definitively identify actively replicating human herpesvirus 5, like viral culture could over 14 days, due to such low representation of viral total nucleic acid overall in the original clinical specimen (0.0005%).

Specimen Type	Patient	Culture	Staining	NGS → TAT: 4 days
Genital Lesion Swab	36yo F	MRC-5 3+ CPE - Day 1	DFA, Human herpesvirus 1 →TAT: 1 day	Human herpesvirus 1 strain H129
Genital Lesion Swab	41yo F	MRC-5 3+ CPE - Day 1	DFA, Human herpesvirus 2 →TAT: 1 day	Human herpesvirus 2 strain SD90e
Nasopharyngeal Swab	3yo F	MRC-5 1+ CPE at Day 14	IFA, Human herpesvirus 5 →TAT: 14 days	Torque Teno virus isolate US32 & Human herpesvirus 5 (trace)

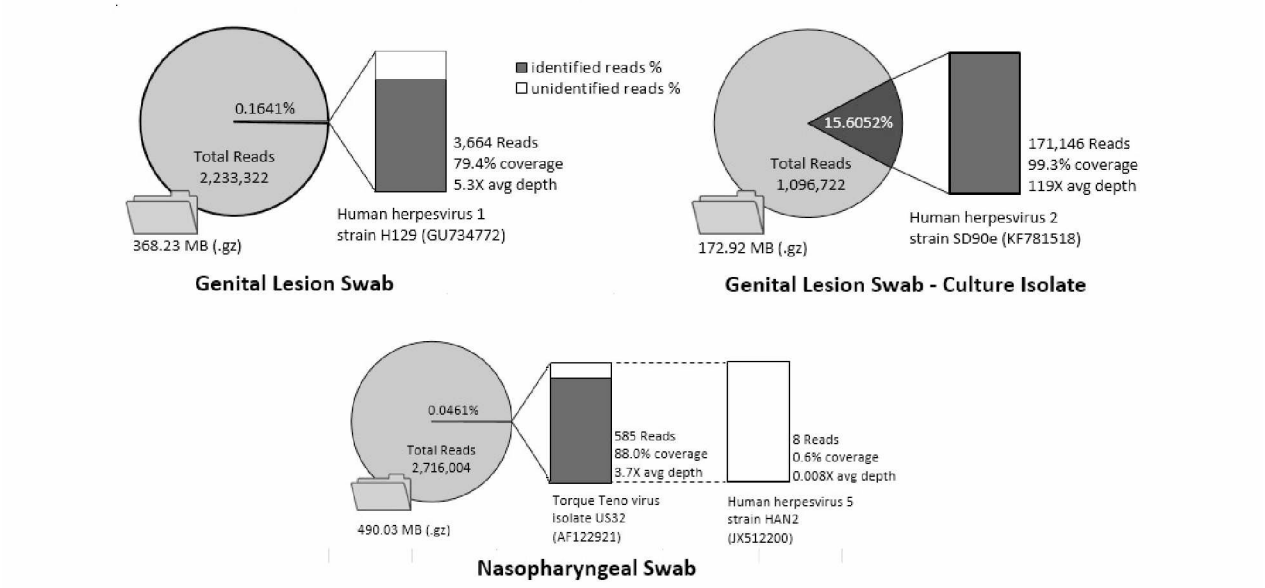


Figure B.2. Comparison of NGS and conventional virology assays for detecting herpesvirus infection.

Further characterization of viral hepatitis C and G

Methodologies to compare NGS's ability to detect hepatitis C was not evaluated, but rather sequence analysis was performed to discover how NGS could be used to further characterize the virus. Specimens were all positive for genotype 1a HCV and specific isolates were identified by PathSeq™Virome (Figure 3).

New direct acting antiviral therapies have been designed to target and impair the functions of non-structural proteins, NS3/4A, NS5A, and NS5B. In response, breakthrough mutations in these particular viral genes have demonstrated antiviral resistance to certain HCV antiviral therapies [29-33]. Figure 3 summarizes resistance data obtained by NGS for each protein targeted as well as the antiviral therapies that are associated. Isolate identified as V60-like was found to have one mutation affecting susceptibility to NS5A-inhibitors (M38V) and 2 mutations affecting susceptibility to N3/NS4-inhibitors (T54S and Q80K). Three other isolates, identified as V179-, V173-, and V269-like, each showed 1 mutation leading to reduced susceptibility to N3/NS4-inhibitors (Q80K).

Figure 3 also shows that hepatitis G virus (HGV) isolates were simultaneously identified in three of the five specimens tested (60%). It is known that HGV infections are closely associated with HCV infections due to parallel routes of transmission [34, 35]. Although not commonly practiced, more thorough surveillance of HGV infections in humans may be necessary due to indications that it is a significant player in determining the course and prognosis of other diseases such as HIV, HCV, and even diseases of the brain [5, 35]. NGS is an appropriate method in which to detect and characterize both HCV and HGV viruses in parallel.

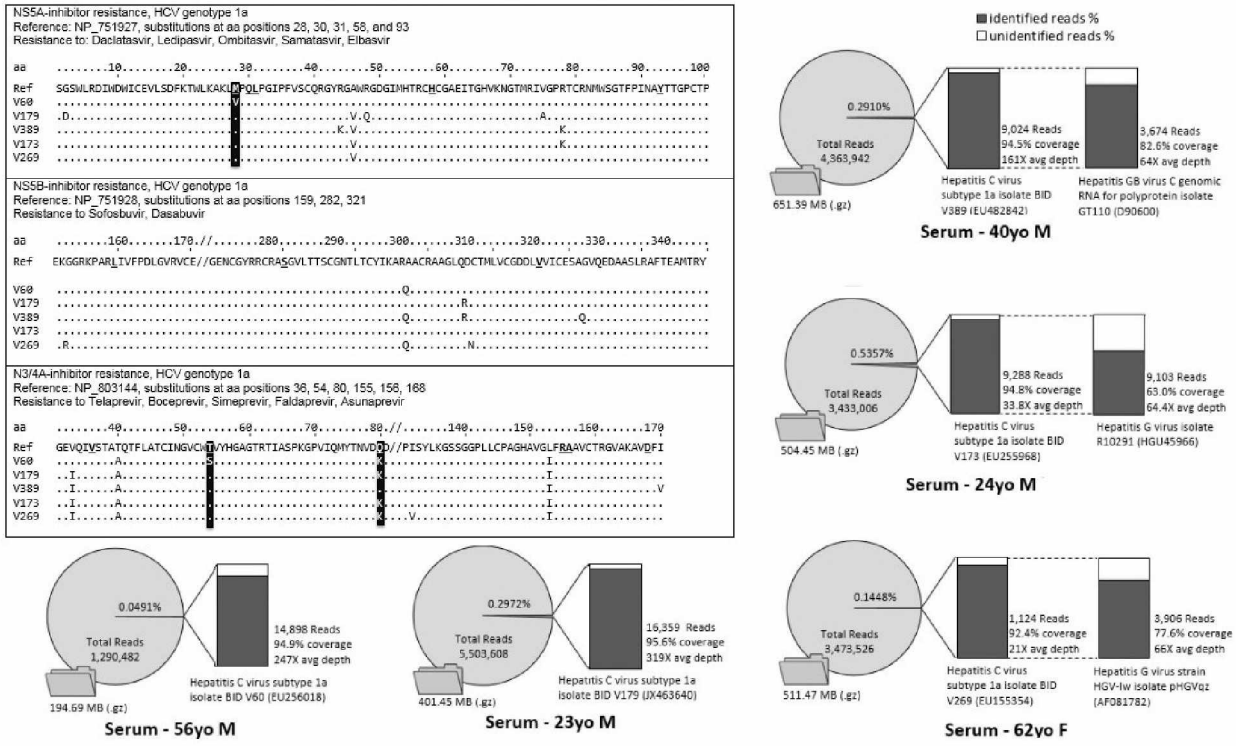


Figure B.3. Detection and characterization of hepatitis C and G viruses in 5 different sera using NGS.
Three groups of antiviral drugs are shown targeting proteins NS5A, NS5B, and N3/N4A. Isolates (i.e. V60, V179) should be thought of as V60-like or V179-like, as identified by sequence analysis. Blackened columns of amino acids indicate that one or more isolates were found to have mutations at those positions associated with antiviral resistance.

Antiviral resistance of influenza viruses in clinical specimens

Data is provided for six nasopharyngeal swab specimens containing influenza viruses tested by PCR methods and NGS (Figure 4). NGS results detected regions of the genome attributed to antiviral drug resistance, as previously described [36-38]. Genome-wide diagrams of alignments to reference genomes produced by Tablet illustrate the various coverages obtained for each clinical specimen (Figure 4). Large variations amongst the results are attributed to the quantity and quality of virus in the original clinical specimen since enrichments techniques such as filtration or centrifugation were not used. Identification of polymorphisms, especially those occurring at specific positions of the neuraminidase gene, may indicate various levels of antiviral resistance to neuraminidase inhibitors and serves as an important piece of information in terms of influenza surveillance. No variants were detected amongst the clinical specimens analyzed (all wild type). For influenza A/H3, one specimen was missing sequence information was for one amino acid motif, 119. For influenza B, one specimen was missing the entire NA gene and could not be analyzed.

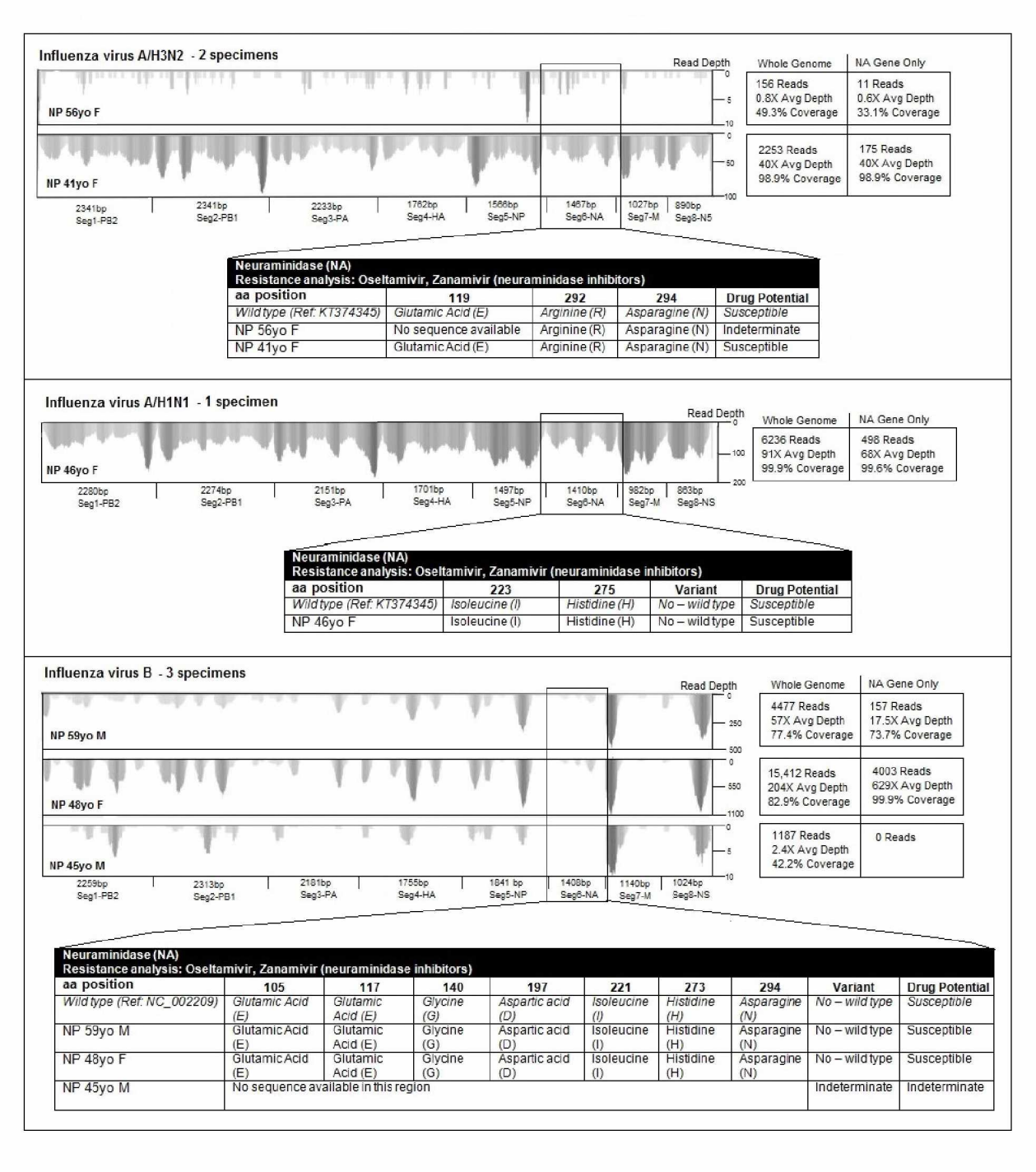


Figure B.4. Antiviral resistance characterization of influenza viruses using NGS Genome-wide views as produced by Tablet. Figure is split into 3 influenza virus types, A/H3N2 viruses, A/H1N1 viruses, and B viruses. Whole genome metrics are compared to those obtained for the NA gene only. Segment 6, the NA gene, is boxed out for each specimen and exploded into a table describing the outcomes of antiviral resistance analysis performed for each type of influenza virus.

Results of a proficiency panel intended for laboratories using pyrosequencing methods are compared to those obtained by NGS (Table 1). Results of NGS had 100% concordance with pyrosequencing results for distinguishing wild-type and variant viruses by identifying mutations in the specific amino acid motifs in the neuraminidase gene as an indicator of antiviral resistance. NGS revealed low coverage sequence reads in negative specimens (10 reads each for PT-A-1 and PT-A-4) revealing the need to establish a standard in which to confidently distinguish non-specific and specific reads in NGS data.

Table B.1
NGS proficiency compared to pyrosequencing methodology for detecting antiviral resistance in influenza A virus

			Pyrosequencing			NGS					
			275		223	275		223	NA Reads	% NA	Depth
H1N1	**Real-Time PCR**	**Variant**	**H%**	**Y%**	**I%**	**H%**	**Y%**	**I**			
PT-A-1	ND	indeterminate	--	--	--	--	--	--	10	35.9%	1.3
PT-A-2	H1N1pdm09	H275	100	0	100	100	0	100	5728	99.6%	850
PT-A-3	H1N1pdm09	H275Y	<10	>90	100	8	92	100	275	98.6%	38.3
PT-A-4	ND	indeterminate							10	48.0%	1.6
PT-A-5	H1N1pdm09	H275Y	60-65	35-40	100	68	42	100	3165	99.8%	436
PT-A-6	H1N1pdm09	H275	100	0	100	100	0	100	2589	98.9%	361
PT-A-7	H1N1pdm09	H275	100	0	100	100	0	100	1364	99.1%	204
PT-A-8	H1N1pdm09	H275Y	60-65	35-40	100	59	41	100	3624	99.7%	497
PT-A-9	H1N1pdm09	H275Y	100	0	100	100	0	100	8878	99.6%	1427
PT-A-10	H1N1pdm09	H275, I223K	100	0	0	100	0	0	4924	99.6%	764
PT-B-1	Not tested	H275	100	0	100	100	0	100	12,788	99.8%	639
PT-B-2	Not tested	H275Y	0	100	100	0	100	100	8,031	99.9%	402
PT-B-3	Not tested	H275, I223K	100	0	0	100	0	0	10,139	99.9%	508
PT-B-4	Not tested	H275Y	60-65	35-40	100	60	40	100	11,496	99.9%	575
H3N2			**E119**	**R292**	**N294**	**E119**	**R292**	**N294**			
PT-B-5	Not tested		100	100	100	100	100	100	8,059	99.9%	418
PT-B-6	Not tested	E119V	0 ($V_{100\%}$)	100	100	0 ($V_{100\%}$)	100	100	12,878	99.9%	670
PT-B-7	Not tested	R292K	100	0 ($K_{100\%}$)	100	100	0 ($K_{100\%}$)	100	30,550	99.9%	1,585

Two separate panels of proficiency specimens (A & B) were tested using NGS and compared to results obtained by laboratories performing pyrosequencing for currently circulating influenza A viruses. Known amino acid positions are unique to each type of neuraminidase gene (N1 or N2) separating out the viruses and varying the positions monitored for antiviral resistance.

Non-specific viral sequencing reads in NGS data

Low coverage viral sequencing is an issue when working with clinical specimens. Collections from patients represent a wide array of pathogen quantities and qualities. Although sequencing is becoming more competitive with other conventional methodologies in terms of cost, there is a need for increasing sequencing depth in order to detect pathogenic viruses in clinical specimens. Deeper sequencing also allows for greater chance for detecting non-specific viral reads. For instance, other viruses were considered "poorly" detected by Pathseq™Virome in each clinical specimen; however, these results could not be substantiated by GSNAP alignments due to very few reads and coverages.

Like other clinical assays, NGS needs a cutoff to determine the true presence of a pathogen versus carry-over or contamination between specimens or other non-specific reads. True negative proficiency specimens (PT-A-1 and PT-A-4) contained 10 reads aligning to the H1N1 NA reference gene (false positive) whereas the human herpesvirus 5 that grew from a nasopharyngeal swab also only had 10 reads. Indeterminate results, such as these, may need to undergo repeat testing where more involved enrichment techniques can be employed to determine the true presence of a virus in low-titer clinical specimens.

Discussion:

We investigated applications of NGS in a clinical laboratory to detect pathogenic viruses in common specimen types and compared NGS data to that which is obtained by more conventional methods. In most cases, with the exception of one, information retrieved by NGS met or exceeded that of conventional methodologies. NGS proves to be a laboratory tool capable

of not only detecting pathogenic viruses in clinical specimens, but also predicting the effects of drug treatment as well.

Through increased use of NGS technologies, reference databases of whole genome sequences can grow and enhance a laboratory's ability to identify sequencing reads. NGS will change our approach as laboratorians and improve our ability to detect and more fully characterize agents of infectious disease in clinical specimens.

Competing Interests: The authors declare no competing interests or conflict of interest.

Ethical Approval: This research project was reviewed and approved by the University of Alaska Fairbanks Institutional Review Board (IRB) (Approval letter No. 667418-2).

Acknowledgements: Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103395 and by an equipment grant of Alaska State Public Health Laboratories. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the NIH and Alaska State Public Health Laboratories. All authors are employees of the State of Alaska's Government, Department of Health and Social Services, Division of Public Health, Section Laboratories. The authors wish to thank the staff at the Alaska State Public Health Virology Laboratory for their technical expertise in the evaluation of the laboratory specimens.

References

1. Köser, C.U., et al., Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. PLoS Pathogens, 2012. 8(8): p. e1002824.

2. Bzhalava, D., et al., Unbiased approach for virus detection in skin lesions. PLoS One, 2013. 8(6): p. e65953.

3. Cheval, J., et al., Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. J Clin Microbiol, 2011. 49(9): p. 3268-75.

4. Chan, B.K., et al., Deep sequencing to identify the causes of viral encephalitis. PLoS One, 2014. 9(4): p. e93993.

5. Kriesel, J.D., et al., Deep sequencing for the detection of virus-like sequences in the brains of patients with multiple sclerosis: detection of GBV-C in human brain. PLoS One, 2012. 7(3): p. e31886.

6. Moore, R.A., et al., The sensitivity of massively parallel sequencing for detecting candidate infectious agents associated with human tissue. PLoS One, 2011. 6(5): p. e19838.

7. Yozwiak, N.L., et al., Virus identification in unknown tropical febrile illness cases using deep sequencing. PLoS Negl Trop Dis, 2012. 6(2): p. e1485.

8. Radford, A.D., et al., Application of next-generation sequencing technologies in virology. Journal of General Virology, 2012. 93(Pt 9): p. 1853-1868.

9. Arroyo, L.S., et al., Next generation sequencing for human papillomavirus genotyping. Journal of Clinical Virology, 2013. 58(2): p. 437-442.

10. Flaherty, P., et al., Ultrasensitive detection of rare mutations using next-generation targeted resequencing. Nucleic Acids Res, 2012. 40(1): p. e2.

11. Meiring, T.L., et al., Next-generation sequencing of cervical DNA detects human papillomavirus types not detected by commercial kits. Virol J, 2012. 9: p. 164.

12. Sijmons, S., M. Van Ranst, and P. Maes, Genomic and functional characteristics of human cytomegalovirus revealed by next-generation sequencing. Viruses, 2014. 6(3): p. 1049-72.

13. Watson, S.J., et al., Viral population analysis and minority-variant detection using short read next-generation sequencing. Philos Trans R Soc Lond B Biol Sci, 2013. 368(1614): p. 20120205.

14. Han, Y., et al., Analysis of hepatitis B virus genotyping and drug resistance gene mutations based on massively parallel sequencing. J Virol Methods, 2013. 193(2): p. 341-7.

15. Messiaen, P., et al., Ultra-deep sequencing of HIV-1 reverse transcriptase before start of an NNRTI-based regimen in treatment-naive patients. Virology, 2012. 426(1): p. 7-11.

16. Sahoo, M.K., et al., Detection of cytomegalovirus drug resistance mutations by next-generation sequencing. J Clin Microbiol, 2013. 51(11): p. 3700-10.

17. Mathonet, P. and C.G. Ullman, The Application of Next Generation Sequencing to the Understanding of Antibody Repertoires. Front Immunol, 2013. 4: p. 265.

18. Lecuit, M. and M. Eloit, The human virome: new tools and concepts. Trends Microbiol, 2013. 21(10): p. 510-5.

19. Yang, J., et al., Unbiased parallel detection of viral pathogens in clinical samples by use of a metagenomic approach. J Clin Microbiol, 2011. 49(10): p. 3463-9.

20. Isakov, O., S. Modai, and N. Shomron, Pathogen detection using short-RNA deep sequencing subtraction and assembly. Bioinformatics, 2011. 27(15): p. 2027-30.

21. Wu, T.D. and S. Nacu, Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics, 2010. 26(7): p. 873-81.

22. Milne, I., et al., Tablet—next generation sequence assembly visualization. Bioinformatics, 2010. 26(3): p. 401-402.

23. Weinberg, G.A., et al., Field evaluation of TaqMan Array Card (TAC) for the simultaneous detection of multiple respiratory viruses in children with acute respiratory infection. J Clin Virol, 2013. 57(3): p. 254-60.

24. Wangchuk, S., et al., Influenza surveillance from November 2008 to 2011; including pandemic influenza A(H1N1)pdm09 in Bhutan. Influenza and Other Respiratory Viruses, 2013. 7(3): p. 426-430.

25. W., D., Adenoviruses Medical Microbiology. 4th edition., 1996(Chapter 67).

26. Ebner, K., W. Pinsker, and T. Lion, Comparative Sequence Analysis of the Hexon Gene in the Entire Spectrum of Human Adenovirus Serotypes: Phylogenetic, Taxonomic, and Clinical Implications. Journal of Virology, 2005. 79(20): p. 12635-12642.

27. Kincaid, R.P., et al., A Human Torque Teno Virus Encodes a MicroRNA That Inhibits Interferon Signaling. PLoS Pathog, 2013. 9(12): p. e1003818.

28. Navarro, D., Expanding role of cytomegalovirus as a human pathogen. Journal of Medical Virology, 2015: p. n/a-n/a.

29. Ahmed, A. and D. Felmlee, Mechanisms of Hepatitis C Viral Resistance to Direct Acting Antivirals. Viruses, 2015. 7(12): p. 2968.

30. Fridell, R.A., et al., Genotypic and phenotypic analysis of variants resistant to hepatitis C virus nonstructural protein 5A replication complex inhibitor BMS-790052 in humans: in vitro and in vivo correlations. Hepatology, 2011. 54(6): p. 1924-35.

31. Bagaglio, S., et al., Frequency of Natural Resistance within NS5a Replication Complex Domain in Hepatitis C Genotypes 1a, 1b: Possible Implication of Subtype-Specific Resistance Selection in Multiple Direct Acting Antivirals Drugs Combination Treatment. Viruses, 2016. 8(4): p. 91.

32. Lemm, J.A., et al., Identification of Hepatitis C Virus NS5A Inhibitors. Journal of Virology, 2010. 84(1): p. 482-491.

33. Ji, H., et al., Next generation sequencing of the hepatitis C virus NS5B gene reveals potential novel S282 drug resistance mutations. Virology, 2015. 477: p. 1-9.

34. Tanaka, E., et al., Past and Present Hepatitis G Virus Infections in Areas Where Hepatitis C is Highly Endemic and Those Where It Is Not Endemic. Journal of Clinical Microbiology, 1998. 36(1): p. 110-114.

35. Reshetnyak, V.I., T.I. Karlovich, and L.U. Ilchenko, Hepatitis G virus. World Journal of Gastroenterology: WJG, 2008. 14(30): p. 4725-4734.

36. Rubaiyea Farrukee, S.-K.L., Jeff Butler, Raphael T.C. Lee, Sebastian Maurer-Stroh, Danielle Tilmanis, Sheena Sullivan, Jennifer Mosse, Ian G. Barr and Aeron C. Hurt, Influenza viruses with B/Yamagata- and B/Victoria-like neuraminidases are differentially affected by mutations that alter antiviral susceptibility. Journal of Antimicrobial Chemotherapy, 2015.

37. Gubareva, L.V., Molecular mechanisms of influenza virus resistance to neuraminidase inhibitors. Virus Res, 2004. 103(1-2): p. 199-203.

38. Ferraris, O. and B. Lina, Mutations of neuraminidase implicated in neuraminidase inhibitors resistance. Journal of Clinical Virology, 2008. 41(1): p. 13-19.

39. Colgrove, R., et al., Genomic sequences of a low passage herpes simplex virus 2 clinical isolate and its plaque-purified derivative strain. Virology, 2014. 450-451: p. 140-5.

Appendix C

Next generation sequencing in clinical virology diagnostics[7]

Abstract

Next generation sequencing (NGS) is a new technology that can be used for broad detection of infectious pathogens and is rapidly becoming an essential platform in clinical laboratories. It is not known how NGS will displace or enhance gold standard methodologies in infectious disease diagnosis at this time, but investigations have begun to understand its potential. Our objective was to investigate the feasibility and application of NGS technology in public health laboratories and compare NGS technology with conventional methods in terms of pathogenic virus detection.

Introduction

Methodologies to detect pathogenic viruses in clinical specimens have transitioned from classic cell culture and antibody-antigen techniques to more sensitive molecular methods such as polymerase chain reaction (PCR). The targeted nature of these methodologies inhibit their ability to accommodate the true diversity of human pathogens in a clinical specimen, especially viruses [1]. Next generation sequencing (NGS) technologies are quickly demonstrating their ability to provide broad detection of infectious agents in a target-independent manner [2-7]. NGS has many advantages beyond the improved detection of all suspected, unsuspected, or even novel pathogens in a clinical specimen [8]. Familiarization with pathogen genomic sequences within clinical specimens enhances our understanding of infectious disease through further discovery of pathogen variability and genotyping [9-11], drug resistance or response to therapy [12], vaccine development and efficacy monitoring [13], and further characterization of the

[7] Original publication as cited: Parker J, Chen, J. *Next generation sequencing in clinical virology diagnostics.* Clinical Lab International, Feb/Mar 2017:6-9.

metagenome [14]. The use of NGS for routine use in clinical diagnostics is emerging with its own set of limitations and challenges [13, 15]. Focusing on viruses of public health importance, we compared the performance of NGS alongside other more common viral detection methodologies.

Conventional methods vs. NGS

We investigated applications of NGS in a clinical laboratory to detect pathogenic viruses in common specimen types and compared NGS data to that which could be obtained by more conventional methods for detecting and characterizing the following viruses of public health importance: adenovirus, herpesvirus, hepatitis C virus, and influenza [16]. We compared results obtained by NGS to viral culture, immunofluorescence staining, serum neutralization, and PCR in terms of turnaround time as well as the clinical relevance of the information obtained.

Table 1 describes the turnaround time of conventional methods to NGS for detecting adenoviruses and herpesviruses, both DNA viruses. The amount of time it takes to grow a virus in culture is variable, ranging from 1 day for herpes simplex viruses to 18 days for adenoviruses. All NGS data could be obtained in 4 days, which includes nucleic acid extraction, sequencing library preparation, sequencing, and data analysis. Although most laboratories are not currently equipped with in-house bioinformaticians, much of the analysis can be done simply using common sequencing analyzing software and the quickly growing number of applications online. For data analysis, we used PathSeq™Virome which enabled us to feed large read files into the application which would generate a report describing the viruses present, including a "detection score" to distinguish strong and weak presence. NGS data provided much more information regarding the exact isolate which may aid health professionals in tracking and relating individual

cases with others. Group C adenoviruses are treatable with cidofovir and NGS data was able to identify the amino acid motif that most affects antiviral resistance.

Hepatitis C virus (HCV) is a growing concern for public health and tends to be difficult to design targeted methodologies around due to the high variability of viral genomes known, even within the same patient. NGS is a powerful tool for characterizing HCV infections and, in our experience, more informational than targeted genotyping assays (Table 2). Since we were able to sequence nearly the entire HCV genome (coverage ranged from 92.4% - 95.6%), data could be generated describing the mutations at key locations across the genomes that are known to cause drug resistance. Antiviral resistance is also critical when characterizing current circulating influenza virus strains and NGS was able to identify viruses that would be considered susceptible to neuraminidase inhibitors (Table 3). In two cases, the viral load of the specimen was too low to achieve good genome coverage across the neuraminidase gene, but this issue could be resolved by screening specimens for high titer (i.e. qPCR) or utilizing enrichment techniques such as ultracentrifugation or filtration of other background nucleic acid.

In most cases, with the exception of one specimen that was unable to definitively identify a cytomegalovirus (HSV5, Table 1), information retrieved by NGS met or exceeded that of conventional methodologies. NGS proves to be a laboratory tool capable of not only detecting pathogenic viruses in clinical specimens, but also predicting the effects of drug treatment as well.

Table C.1: Comparison of methods for detecting DNA viruses

Specimen	Conventional Methodologies				Next-generation sequencing	
	Viral Culture	IFA	Serum Neutralization	TAT	PathSeq™Virome	TAT
NP swab	3+ CPE on day 2	Positive Adenovirus	Adenovirus (type 2)	14	Human adenovirus C strain human/ARG/A15932/2002/2 (Cidofovir susceptible)	4
Nasal swab	4+ CPE on day 8	Positive Adenovirus	Adenovirus (type 3)	18	Human adenovirus B strain human/USA/UFL_Adv3a51/2007/3	4
Genital swab	3+ CPE on day 1	Positive HSV1	Not performed	1	Human herpesvirus 1 strain H129	4
Genital swab	3+ CPE on day 1	Positive HSV2	Not performed	1	Human herpesvirus 2 strain SD90e	4
NP Swab	1+ CPE on day 14	Positive HSV5	Not performed	14	Torque Teno virus isolate US32 & Human herpesvirus 5 (trace)	4

TAT = turnaround time

Table C.2: NGS for characterizing hepatitis C viruses and potential resistance to infection inhibitors

Specimen	PCR	PathSeq™Virome	HCV Drug Resistance Markers*		
			NS5A-inhibitor	NS5B-inhibitor	N3/4A-inhibitor
Serum	Positive HCV	Hepatitis C virus subtype 1a isolate BID V60	1	susceptible	2
Serum	Positive HCV	Hepatitis C virus subtype 1a isolate BID V179	susceptible	susceptible	1
Serum	Positive HCV	Hepatitis GB virus C genomic RNA for polyprotein isolate GT110; Hepatitis C virus subtype 1a isoate BID V389	susceptible	susceptible	susceptible
Serum	Positive HCV	G virus strain HGV-lw isolate pHGVqz; Hepatitis C virus subtype 1a isolate BID V269	susceptible	susceptible	1
Serum	Positive HCV	Hepatitis G virus isolate R10291; Hepatitis C virus subtype 1a isolate BID V173	susceptible	susceptible	1

*Number of mutations within these particular genes that demonstrate resistance to various HCV antiviral therapies

Table C.3: NGS for characterizing influenza viruses

Specimen	**NGS**	**Neuraminidase Inhibitors – Drug Resistance Markers**
NP swab	Influenza A/H3	Amino acid positions: 110, 292, and 294 – Indeterminate
NP swab	Influenza A/H3	Amino acid positions: 110, 292, and 294 – Susceptible (wild type)
NP swab	Influenza A/H1N1pdm09	Amino acid positions: 223 and 275 – Susceptible (wild type)
NP swab	Influenza B (Yamagata Lineage)	Amino acid positions: 105, 117, 140, 197, 221, 273, 294 – Susceptible (wild type)
NP swab	Influenza B (Yamagata Lineage)	Amino acid positions: 105, 117, 140, 197, 221, 273, 294 – Susceptible (wild type)
NP swab	Influenza B (Victoria Lineage)	Amino acid positions: 105, 117, 140, 197, 221, 273, 294 – Susceptible (wild type)

Summary

Through increased use of NGS technologies, reference databases of whole genome sequences can grow and enhance our ability to more definitively identify sequencing reads. Although this review describes conventional methods vs. NGS for detecting specific viruses, there was also evidence of the presence of co-infecting viruses such as hepatitis G and Torque Teno virus that weren't originally targeted. The standard 4 day turnaround time needed to complete NGS could be improved with extraction and library preparation automation, as well as advances in sequencing technology (each run ~40 hours). Based on our laboratory's experience and the growing body of literature, NGS will change our approach as clinical laboratorians and improve our ability to detect and more fully characterize agents of infectious disease in clinical specimens in a non-targeted manner.

References

1. Köser, C.U., et al., Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. PLoS Pathogens, 2012. 8(8): p. e1002824.
2. Bzhalava, D., et al., Unbiased approach for virus detection in skin lesions. PLoS One, 2013. 8(6): p. e65953.
3. Cheval, J., et al., Evaluation of high-throughput sequencing for identifying known and unknown viruses in biological samples. J Clin Microbiol, 2011. 49(9): p. 3268-75.
4. Chan, B.K., et al., Deep sequencing to identify the causes of viral encephalitis. PLoS One, 2014. 9(4): p. e93993.
5. Kriesel, J.D., et al., Deep sequencing for the detection of virus-like sequences in the brains of patients with multiple sclerosis: detection of GBV-C in human brain. PLoS One, 2012. 7(3): p. e31886.
6. Moore, R.A., et al., The sensitivity of massively parallel sequencing for detecting candidate infectious agents associated with human tissue. PLoS One, 2011. 6(5): p. e19838.
7. Yozwiak, N.L., et al., Virus identification in unknown tropical febrile illness cases using deep sequencing. PLoS Negl Trop Dis, 2012. 6(2): p. e1485.

8. Radford, A.D., et al., Application of next-generation sequencing technologies in virology. Journal of General Virology, 2012. 93(Pt 9): p. 1853-1868.

9. Arroyo, L.S., et al., Next generation sequencing for human papillomavirus genotyping. Journal of Clinical Virology, 2013. 58(2): p. 437-442.

10. Flaherty, P., et al., Ultrasensitive detection of rare mutations using next-generation targeted resequencing. Nucleic Acids Res, 2012. 40(1): p. e2.

11. Meiring, T.L., et al., Next-generation sequencing of cervical DNA detects human papillomavirus types not detected by commercial kits. Virol J, 2012. 9: p. 164.

12. Sijmons, S., M. Van Ranst, and P. Maes, Genomic and functional characteristics of human cytomegalovirus revealed by next-generation sequencing. Viruses, 2014. 6(3): p. 1049-72.

13. Watson, S.J., et al., Viral population analysis and minority-variant detection using short read next-generation sequencing. Philos Trans R Soc Lond B Biol Sci, 2013. 368(1614): p. 20120205.

14. Han, Y., et al., Analysis of hepatitis B virus genotyping and drug resistance gene mutations based on massively parallel sequencing. J Virol Methods, 2013. 193(2): p. 341-7.

15. Messiaen, P., et al., Ultra-deep sequencing of HIV-1 reverse transcriptase before start of an NNRTI-based regimen in treatment-naive patients. Virology, 2012. 426(1): p. 7-11.

16. Parker, J. and J. Chen, Application of next generation sequencing for the detection of human viral pathogens in clinical specimens. J Clin Virol, 2017. 86: p. 20-26.